General Meeting
of the
NASA Astrobiology Institute

Chemistry and Origins of Living Systems

Could Biochemistry Have Hydrothermal Origins?

George D. Cody
Geophysical Laboratory
Carnegie Institution of Washington
5251 Broad Branch Rd., NW, Washington, DC
USA
Cody@GL.CIW.EDU

Nabil Boctor
Geophysical Laboratory
Carnegie Institution of Washington
5251 Broad Branch Rd., NW, Washington, DC
USA

Robert M. Hazen
Geophysical Laboratory
Carnegie Institution of Washington
5251 Broad Branch Rd., NW, Washington, DC
USA

James Scott
Geophysical Laboratory
Carnegie Institution of Washington
5251 Broad Branch Rd., NW, Washington, DC
USA

Anurag Sharma
Geophysical Laboratory
Carnegie Institution of Washington
5251 Broad Branch Rd., NW, Washington, DC
USA

Hatten S. Yoder, Jr.
Geophysical Laboratory
Carnegie Institution of Washington
5251 Broad Branch Rd., NW, Washington, DC
USA

The proposal that deep-sea hydrothermal vent systems might be an ideal environment for the emergence of Earth's first life is particularly compelling in that it greatly expands the possibility for life beyond the narrow band of solar luminosity that supports surface liquid water; thus, this theory accommodates the possibility of biochemistry, similar to that of terrestrial life, emerging within other planetary bodies in this and other solar systems. However, the theory is poorly constrained due to a lack of experimental data.

In order to remedy this deficiency we have run high pressure hydrothermal reactions using sealed gold tube reactors and an internally heated high pressure apparatus. First, we studied the hydrothermal reactions of pure pyruvic acid and citric acid systems (sans catalysts), respectively, in CO_2 - H_2 bearing aqueous fluids in the temperature range of 150-350 °C and 0.5 to 5.0 Kbar. At this stage our principal focus is in identifying the

effect of pressure on reaction selectivity in systems that exhibit multi-channel reactions. Towards this end, we have focused on two relatively simple systems in order to gain fundamental information on the kinetics of reactions that typify biological processes, albeit under strictly abiotic conditions. The data derived will provide a foundation for subsequent work on the potential for abiotic synthesis of classically bio-organic compounds at extremes of temperature and pressure, conditions that typify deep ocean hydrothermal vents.

In the pyruvic acid system we set out to determine whether pressure would favor the synthesis of oxaloacetic acid through an electrophilic addition of CO_2 at elevated temperatures. We observe three reaction channels operating in series and parallel. One of these channels yields appreciable quantities of methyl succinic acid, that forms through an Aldol condensation followed by a decarboxlyation. A second reaction channel involves the straightforward decarboxylation of pyruvic acid to form acetic acid, CO_2, and H_2. The third reaction channel produces a very interesting and complex suite of compounds that exhibit amphiphilic qualities. This particular reaction channel operates, in general, through sequential Aldol condensations, Diels-Alder cycloadditions, decarboxylations, dehydrations, dehydrogenations, and/or hydrogenations. Significant and systematic pressure and temperature induced changes in the distribution of molecules within this suite are observed.

The system citric acid in CO_2 - H_2 bearing aqueous fluids is equally interesting. As would be expected, the principal reactions involve mono, double, and triple decarboxylations; with and without dehydration, as well as hydrogenation. The product distributions exhibit strong temperature and pressure selectivity. In general the citric system under the range of conditions explored exhibits catabolic chemistry. This is due, principally the apparent irreversibility of acid catalyzed decarboxylations. Surprisingly, pressure enhances the kinetics of these reactions, thus greatly accelerates the catabolic evolution of the system. The citric acid system, does however, serve as a useful and relatively simple system to explore the role of pressure in essentially pure ionic aqueous organic chemistry.

We have further set out to explore whether the catalytic capabilities of naturally occurring transition metal sulfide minerals could be used to promote reactions that might serve biochemical function. The role that transition metal sulfides play for extant life, within the active centers of enzymes, is extremely broad. We have focused first on mineral sulfides to promote carbon insertion reactions; e.g. to mimic the activity of the primitive CO monohydrogenase enzyme found in some hyperthermophilic autotrophs. Our results show that while Ni and Co monosulfides operated as particularly good catalysts, a wide range of minerals including Fe and Cu sulfides can promote the carbonyl insertion reactions. These results lend support to idea of hydrothermal vent systems may have served as primordial proto-biochemical reactors.

A Chiroselective Peptide Replicator and its Relevance to Issues Concerning the Origin of Homochirality on Earth

M. Reza Ghadiri
Departments of Chemistry and Molecular Biology, and The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 N. Torrey Pines Rd. La Jolla, CA 92037
USA
Ghadiri@scripps.edu

Alan Saghatelian
Departments of Chemistry and Molecular Biology, and The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 N. Torrey Pines Rd. La Jolla, CA 92037
USA
saghat@scripps.edu

Yohei Yokobayashi
Departments of Chemistry and Molecular Biology, and The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 N. Torrey Pines Rd. La Jolla, CA 92037
USA
yohei@scripps.edu

Kathy Soltani
Departments of Chemistry and Molecular Biology, and The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 N. Torrey Pines Rd. La Jolla, CA 92037
USA

The origin of homochirality in living systems is often attributed to the generation of enantiomeric differences in a pool of chiral prebiotic molecules, but none of the possible physiochemical processes considered can produce the significant imbalance required if homochiral biopolymers are to result from simple coupling of suitable precursor molecules. This implies a central role either for additional processes that can selectively amplify an initially minute enantiomeric difference in the starting material, or for a nonenzymatic process by which biopolymers undergo chiroselective molecular replication. Given that molecular self-replication and the capacity for selection are necessary conditions for the emergence of life, chiroselective replication of biopolymers seems a particularly attractive process for explaining homochirality in nature. Here we report that a 32-residue peptide replicator, designed according to our earlier principles, is capable of efficiently amplifying homochiral products from a racemic mixture of peptide fragments through a chiroselective autocatalytic cycle. The chiroselective amplification process discriminates between structures possessing even single stereochemical mutations within otherwise homochiral sequences. Moreover, the system exhibits a dynamic stereochemical "editing" function making use of heterochiral sequences that arise through uncatalysed background reactions to catalyse the production of the

homochiral product. These results support the idea that self-replicating polypeptides could have played a key role in the origin of homochirality on Earth.

Organic Synthesis in Simulated Interstellar Ice Analogs

Jason P. Dworkin
Astrophysics Branch
SETI Institute and NASA – Ames Research Center
Mail Stop 245-6, Moffett Field, CA 94035
USA
jdworkin@mail.arc.nasa.gov

Max P. Bernstein
Astrophysics Branch
SETI Institute and NASA – Ames Research Center
Mail Stop 245-6, Moffett Field, CA 94035
USA

Scott A. Sandford
Astrophysics Branch
NASA – Ames Research Center
Mail Stop 245-6, Moffett Field, CA 94035
USA

Louis J. Allamandola
Astrophysics Branch
NASA – Ames Research Center
Mail Stop 245-6, Moffett Field, CA 94035
USA

David W. Deamer
Department of Chemistry and Biochemistry
University of California, Santa Cruz
Santa Cruz, CA 95064
USA

Jamie Elsila
Department of Chemistry
Stanford University
Stanford, CA 94305-5080
USA

Richard N. Zare
Department of Chemistry
Stanford University
Stanford, CA 94305-5080
USA

Comets and carbonaceous micrometeorites may have been significant sources of organic compounds on the early Earth. Ices on grains in interstellar dense molecular clouds contain a variety of simple molecules as well as aromatic molecules of various sizes.

While in these clouds the icy grains are processed by ultraviolet light and cosmic radiation which produces more complex organic molecules.

We have run laboratory simulations to identify the types of molecules which could have been generated photolytically in pre-cometary ices. Experiments were conducted by forming various realistic interstellar mixed-molecular ices with and without polycyclic aromatic hydrocarbons (PAHs) at ~10 K under high vacuum irradiated with UV light from a hydrogen plasma lamp. The residue that remained after warming to room temperature was analyzed by HPLC, and by laser desorption mass spectrometry. The residue contains several classes of compounds which may be of prebiotic significance.

The Origin of Organic Matter in the Solar System: Evidence from Interplanetary Dust Particles

G. J. Flynn
Dept. of Physics
SUNY-Plattsburgh
Plattsburgh, NY 12901 USA
george.flynn@plattsburgh.edu

L. P. Keller
NASA Johnson Space Center
Houston TX 77058 USA

C. Jacobsen and S. Wirick
Dept. of Physics
SUNY-Stony Brook
Stony Brook, NY 11794 USA

The origin of the organic matter in interplanetary materials has not been established. A variety of mechanisms have been proposed, with two extreme cases being a Fisher-Tropsch type process operating in the gas phase of the solar nebula or a Miller-Urey type process, which requires interaction with an aqueous fluid, presumably occurring on an asteroid [1]. In the Fisher-Tropsch case, we might expect similar organic matter in hydrated and anhydrous interplanetary materials. However, aqueous alteration is required in the case of the Miller-Urey process, and we would expect to see organic matter preferentially in interplanetary materials that exhibit evidence of aqueous activity, such as the presence of hydrated silicates.

The types and abundance of organic matter in meteorites have been used as an indicator of the origin of organic matter in the Solar System. Indigenous complex organic matter, including amino acids, has been found in hydrated carbonaceous chondrite meteorites, such as Murchison [1]. Much lower amounts of complex organic matter, possibly only terrestrial contamination, have been found in anhydrous carbonaceous chondrite meteorites, such as Allende, that contain most of their carbon in elemental form [1]. These results seem to favor production of the bulk of the organic matter in the Solar System by aqueous processing on parent bodies such as asteroids, a Miller-Urey process.

However, the hydrated carbonaceous chondrite meteorites have approximately solar abundances of the moderately volatile elements, while all anhydrous carbonaceous chondrite meteorites have significantly lower contents of these moderately volatile elements. Two mechanisms, incomplete condensation or evaporation, both of which involve processing at ~1200^{o} C, have been suggested to explain the lower content of the moderately volatile elements in all anhydrous meteorites. The high temperatures associated with either the incomplete condensation or the subsequent vaporization are sufficient to remove or destroy most volatile organic matter in the anhydrous meteorites, *indicating that the meteorite studies do not constrain the origin of organic matter.*

Most anhydrous interplanetary dust particles (IDPs), dust from comets and asteroids collected from the Earth's stratosphere, have volatile contents higher than the hydrated carbonaceous meteorites, suggesting these IDPs experienced minimal thermal processing. Unequilibrated mineralogy indicates many IDPs never experienced significant heating. Some IDPs have regions with high D/H within a few μm of other phases with lower D/H, indicating these IDPs were not heated to the point of D equilibration since IDP formation.

To infer the origin of organic matter in the Solar System we measured the abundance and types of carbon in *all three kinds* of primitive (not thermally processed) extraterrestrial materials available for laboratory analysis: anhydrous IDPs, hydrated IDPs, and hydrated carbonaceous chondrite meteorites. Because of the small sizes (~10 μm) of the IDPs, the organic carbon in IDPs cannot be characterized by the traditional analysis techniques. We employ two synchrotron-based instruments at the National Synchrotron Light Source:

1) A Scanning Transmission X-ray Microscope (STXM), to map the carbon distribution and perform Carbon X-ray Absorption Near Edge Structure (XANES) spectroscopy.
2) A Fourier Transform Infrared (FTIR) spectrometer, to detect organic compounds.

The analysis methods and results are described in detail in Flynn et al. [2]. STXM carbon maps of 12 IDPs indicate they have carbon contents ranging from a few to 90 vol.-%. The carbon contents are not substantially different in the hydrated and anhydrous subgroups. The 3 hydrated IDPs and 7 of the 9 anhydrous IDPs we examined using the STXM have essentially identical C-XANES spectra, and these spectra are generally similar to the C-XANES spectrum of acid-insoluble organic matter extracted from Murchison, a hydrated carbonaceous chondrite meteorite. Each of these C-XANES spectra shows a large absorption at ~288.5 eV, characteristic of the C=O bond (and distinct from the ~290 eV absorption of the C-O in carbonate), indicating a substantial fraction of the carbon is organic rather than elemental or mineralogical [2]. The IDPs are dominated by anhydrous and hydrated silicates, carbonates, sulfides, and oxides, which have strong absorptions over most of the 4000 to 650 cm^{-1} range we analyze by FTIR. This interferes with the detection of organic features over much of the IR spectrum. These phases do not interfere from 2700 to 3100 cm^{-1}, where C-H stretching vibrations occur. Eight of the 11 IDPs examined by FTIR showed strong $C-H_2$ and $C-H_3$ stretching vibrations of aliphatic hydrocarbon. The 2700 to 3100 cm^{-1} region of the spectra of hydrated and of anhydrous IDPs were indistinguishable, and were very similar to the spectrum of bulk Murchison.

The similarity of the C abundance, C-XANES and FTIR spectra of anhydrous IDPs, hydrated IDPs, and hydrated carbonaceous meteorites are *consistent with the production of much of the organic matter prior to incorporation into the asteroids,* possibly by irradiation of C-bearing ices or by a Fisher-Tropsch type process operating in the gas phase of the nebula or in the interstellar medium (given large D excess in some IDPs).

References: 1) J. R. Cronin, S. Pizzarello, and D. P. Cruikshank (1988) "Organic Matter in Carbonaceous Chondrites…" in *Meteorites and the Early Solar System,* U. of Arizona Press, 819-857. 2) G. J. Flynn, L. P. Keller, C. Jacobsen, S. Wirick, and M. A. Miller

(2000) “Organic Carbon in Interplanetary Dust Particles” in *A New Era in Bioastronomy*, ASP Conference Series, Vol. 213, Astron. Soc. of the Pacific Press, 191-194.

Selective Adsorption of L- and D-Amino Acids on Calcite: Implications for the Origin of Biochemical Homochirality

Robert M. Hazen
Geophysical Laboratory and NASA Astrobiology Institute
Carnegie Institution of Washington
5251 Broad Branch Road NW
Washington, DC 20015 USA
hazen@gl.ciw.edu

Timothy R. Filley
Department of Earth and Atmospheric Science
Purdue University
West Lafayette, IN 47907 USA

Glenn A. Goodfriend
Department of Earth and Environmental Science
George Washington University
Washington, DC 20052 USA

One of life's most distinctive biochemical signatures is its strong selectivity for chiral molecular species, notably L-amino acids and D-sugars. Prebiotic synthesis reactions, with the possible exception of some interstellar processes, yield essentially equal amounts of L- and D-enantiomers. A significant challenge in origin-of-life research, therefore, is to identify natural mechanisms for the homochiral selection, concentration and polymerization of molecules from an initially racemic mixture. Symmetry breaking on a chirally-selective mineral surface offers a viable scenario for the origin of life; according to Lahav (*Biogenesis*. NY: Oxford University Press, 1999, p.259), "if a selective adsorption of chiral amino acids … on certain crystal faces were observed, then the problem of biological homochirality would be possible to comprehend."

We demonstrate that crystals of the common rock-forming mineral calcite ($CaCO_3$), when immersed in a racemic aspartic acid solution, display significant adsorption and chiral selectivity of D- and L-enantiomers on pairs of mirror-related crystal growth surfaces. This selective adsorption is greater on crystals with terraced surface textures, which suggests that D- and L-aspartic acid concentrate along step-like, linear growth features. Selective adsorption of D- and L-amino acids on calcite is thus a plausible geochemical mechanism for the chiral selection and subsequent homochiral polymerization of amino acids on the prebiotic Earth.

Effectiveness of Hydrogen Sulfide as a Reductant in Hydrothermal Systems: Implication for Prebiotic Synthesis of C-H-O-N Compounds

Martin Schoonen
Department of Geosciences
State University of New York at Stony Brook
ESS 220, Stony Brook, NY 11794-2100
USA.
Penn State Astrobiology Research Center
435 Deike Building
Penn State University
University Park, PA 16802
mschoonen@notes.cc.sunysb.edu

Equilibrium thermodynamic calculations indicate that the chemical disequilibrium in vent solutions discharging from submarine hydrothermal systems could have provided the necessary driving force to synthesize metastable C-H-O-N compounds [1, 2]. Carbon in vent solutions discharging from modern systems has an oxidation state of IV (i.e., CO_2-carbon). Unless the overall oxidation state of the Earth's crust was significantly lower, the carbon discharging from ancient vent system would also have been in the IV-state [1]. Barring a very reducing Earth's crust, nitrogen would have discharged as zerovalent N_2 in ancient systems as it does in modern vent systems. Because the redox state of C and N in C-H-O-N compounds is less than IV and 0, respectively, a reductant is needed to make the synthesis possible. While solution composition in primordial vent systems may have been different, there are only a few reductants available in any significant abundance. These are ferrous iron, hydrogen, and hydrogen sulfide. Dissolved ferrous iron is a relatively weak reductant. Furthermore, it is readily sequestered in Fe-S phases. This leaves hydrogen and hydrogen sulfide as possible reductants. The abundance of each of these reductants is governed by the mineral assemblage that effectively controls the redox state of the vent solution as it makes its way toward the point of discharge [1]. Depending on these constraints, either hydrogen or hydrogen sulfide may be the more abundant reductant. Abundance is, however, less important than reactivity.

A survey of the catalysis literature indicates that hydrogen is very reactive if it is present as atomic hydrogen, but largely unreactive as molecular hydrogen. In fact, dissociating molecular hydrogen on a suitable surface is a key step in hydrogenation reactions. Suitable catalysts include various metal surfaces; some limited H_2 dissociation is observed on iron oxides. By contrast, H_2 does not dissociate on pyrite [3]. It is not clear then if any of the mineral constituents commonly present in vent systems (i.e., metal sulfides and metal oxides) can activate some of the discharging H_2. The presence of hydrogen sulfide, often referred to as a poison in the catalysis literature, may also limit the capacity of any mineral surface to activate H_2. If hydrogen is largely ineffective as a reductant then hydrogen sulfide emerges as the single most important reductant.

Hydrogen sulfide is a strong reductant because of the stability of pyrite and elemental sulfur. While thermodynamic calculations show that coupling of H_2S oxidation to the reduction of either N_2 and CO_2 is energetically favorable, there have been only a few experimental studies that have explored these reactions. Furthermore, there are few studies that address the interaction of hydrogen sulfide with (mineral) surfaces that may act as catalysts. Recently, our lab has shown that di-nitrogen can be reduced by hydrogen sulfide to ammonia in the presence of freshly precipitated FeS at a temperature of 120C and a partial pressure of N_2 of 50 bar (Table 1). The yields of these experiments are consistent with a reaction in which hydrogen sulfide is converted to elemental sulfur, not pyrite. To make ammonia the hydrogen sulfide has to dissociate. Modern surface science studies indicate that hydrogen sulfide can dissociate and form atomic hydrogen on the surface of pyrite [3]. We speculate that a similar reaction on the surface of the FeS in our experimental system may have produced the observed ammonia. A similar surface mechanism may have led to the thiol formation in the experimental study of CO_2 reduction by H_2S in the presence of FeS by Heinen and Lauwers [4].

Table 1. Formation of NH_4^+ via N_2 reduction by H_2S (concentrations in micromol/L)

Exp	Starting $[NH_4^+]$	Final $[NH_4^+]$	T, °C	phase	solid
NRTi03	2.2 ±0.2	2.3 ± 0.2	120	L	none
NRSS01	1.6 ±0.2	16.2 ± 0.5	120	L	none
NRTi05	2.4 ±0.2	8.2 ± 0.4	120	L	FeS
NRTi06	3.5 ±0.2	14.1 ± 0.3	120	L	FeS
NRTi09	1.6 ±0.2	12.4 ± 0.4	110	L	FeS
NRTi01	1.9 ±0.2	8.1 ± 0.5	120	LV	FeS
NRTi02	2.0 ±0.2	8.1 ± 0.4	120	LV	none

NH_4^+ concentrations reported here represent an average total dissolved ammonia (NH_3 + NH_4^+) based on six to ten analyses on each fluid. Experiments conducted in Ti-17 (NRTixx) and SS-316 (NRSSxx) tube reactors. Phase: L = liquid, V=vapor.

The study of reactions with H_2S as reductant in hydrothermal systems is complicated by its reactivity toward various metals commonly used as container materials. For example, we found that ammonia was formed in the absence of FeS in 316 stainless steel tubular vessels. Both experiments with entirely liquid-filled 316 ss tubes and tubes with a head space produced ammonia consistent with a conversion of H_2S to S(0). By contrast, experiments in Ti-17 tubes without FeS produced only ammonia if there was a headspace. These results suggest that: 1) an Fe-S phase formed on the surface of the 316ss which promoted the reaction, presumably via a mechanism similar to the one outlined for FeS; and 2) a gas phase reaction on the Ti-17 surface produced ammonia. Ammonia formation via a gas phase reaction on Ti-17 indicates that gaseous H_2S molecules dissociate on the material and combine with co-adsorbed N_2. Dissociation of gaseous H_2S on Ti-17 forming atomic hydrogen is perhaps the critical step. Experimentalists working with Ti-17 vessels should be aware of this gas-phase process as new studies with hydrogen are undertaken.

1 E.L. Shock, Geochemical constraints on the origin of life of organic compounds in hydrothermal systems, Orig. Life Evol. Biosphere 20, 331-367, 1990.
2 M.A.A. Schoonen, Y. Xu and J. Bebie, Energetics and kinetics of the prebiotic synthesis of simple organic acids and amino acids with the FeS/FeS_2 redox couple as reductant, Orig. Life Evol. Biosphere 29, 5-32, 1999.

3 J.M. Guevremont, D.R. Strongin and M.A.A. Schoonen, Thermal Chemistry of H_2S and H_2O on the striated (100) Plane of Pyrite: Unique Reactivity of Defect Sites, American Mineralogist 83, 1246-1255, 1998.
4 W. Heinen and A.-M. Lauwers, Organic sulfur compounds resulting from the interaction of iron sulfide, hydrogen sulfide and carbon dioxide in an anaerobic environment, Orig. Life Evol. Biosphere 26, 131-150, 1996.

The Origins of Evolution

Jack W. Szostak
Department of Molecular Biology
Massachusetts General Hospital
Boston, MA 02114
USA
szostak@molbio.mgh.harvard.edu

David P. Bartel
Whitehead Institute for Biomedical Research
MIT
Cambridge, MA 02142
USA

P. Luigi Luisi
ETH-Zentrum
Institut für Polymere
Universitätstrasse 6, CH-8092 Zürich
Switzerland

The evolution of modern biochemistry is most easily explained by positing a stage in the early evolution of life in which biochemical functions such as information transfer and replication were largely carried out by RNA molecules. The recently determined structure of the ribosome makes this point most dramatically, since it is RNA that is responsible for the synthesis of proteins. The importance of a membrane in providing a compartment boundary has also been stressed by many authors from varying points of view, including the establishment of cellular identity as distinct from the environment, as a locus for metabolism, and as a requisite for energy generation. Here we discuss the role of the membrane compartment in enabling the Darwinian evolution of its enclosed genetic material. We also show that the evolution of a genetically encoded function that enhances compartment replication or survival is a requirement for Darwinian evolution at the cellular level.

The role of a membrane compartment in enabling Darwinian evolution of the genetic material (RNA or otherwise) is most easily understood by imagining what would happen in the absence of a compartment. For the sake of argument, let us imagine an RNA based system, in which RNA molecules with RNA polymerase activity replicate each other. If such RNAs are in free solution, mutants will arise during replication, and some of these may confer properties such as enhanced fidelity or rate of replication. However, such mutant molecules will not have any selective advantage over their parental sequence, or indeed over parasitic inactive molecules, since any replicase molecule can only be replicated by the action of other replicases. In contrast, a system of replicating compartments provides a simple mechanism for keeping molecules that are related by descent in physical proximity with each other. After a period of random segregation, mutant replicases will be in a compartment populated by clonal descendants of the

original mutant sequence. In such a situation, more accurate replicases will replicate each other more accurately, and a selective advantage will be obtained. Furthermore inactive sequences will be segregated away from the rest of the population and will die out, instead of acting as parasites. The replication of a genetic polymer (whether RNA or a simpler progenitor) inside a population of replicating vesicles will therefore lead to the evolution of an increasingly sophisticated replication biochemistry.

However, a vesicle that contains a superior replicase does receive any advantage with respect to its own replication, i.e. evolution is still not operating at the cellular level. For evolution to act at this level, there must be some coupling between the genetic material and the growth and/or survival of the whole cell. Such a coupling could evolve in numerous ways. Perhaps the most straightforward would be the evolution of some new genetically encoded function (e.g. a ribozyme) that catalyzes the synthesis of new amphiphilic molecules for incorporation into the membrane. Another possibility would be the evolution of new genetically encoded structural molecules that facilitate the division of the membrane compartment into daughter vesicles at the appropriate time. There are undoubtedly many other possible mechanisms by which encoded functions could facilitate cellular growth and division. The first time that such a function evolved, Darwinian evolution at the cellular level would have become both possible and inescapable - leading directly to evolution of more and more complex cells.

Atmospheric Aerosols in Prebiotic Chemistry

Adrian F. Tuck
NOAA Aeronomy Laboratory
Boulder, CO 80303
USA
tuck@al.noaa.gov

Veronica Vaida
Department of Chemistry and Biochemistry
University of Colorado
Boulder, CO 80309
USA

It is argued that atmospheric aerosols would have had many advantages as a reaction medium in prebiotic chemistry. This theme, published last October 24 in PNAS, will be reviewed and developed further as regards the generation of amphiphilic molecules, and the chemistry which could have occurred near the tropopause as stratospheric aerosols containing meteoric material coagulated with tropospheric aerosols containing marine material.

References:
Murphy, Thomson & Mahoney, Science 282 1664-1669 (1998)
Ellison, Tuck & Vaida, J Geophys Res 104 11633-11641 (1999)
Dobson, Ellison, Tuck & Vaida, Proc. Nat. Acad. Sci. 97 11864-11868 (2000).

Stability and Transformations of Nitrogenous Compounds Under Hydrothermal Conditions

Jay A. Brandes
Department of Marine Sciences
The University of Texas at Austin
750 Channel View Dr., Port Aransas, TX 78373
USA
brandes@utmsi.utexas.edu

George D. Cody
Robert M. Hazen
Hatten S. Yoder, Jr
Geophysical Laboratory
Carnegie Institution of Washington
USA

Hydrothermal systems have been hypothesized as locations for the genesis of life. However, few experiments have been conducted that examine the interplay of biologically important compounds with minerals associated with these systems. One element generally neglected in studies of prebiotic chemistry is nitrogen. Several key questions in the study of the early Earth's nitrogen cycle are: How and under what conditions was reduced nitrogen formed? How was this reduced nitrogen incorporated into organic compounds, and what were the important parameters controlling the stability of these nitrogen containing organic compounds? A series of high pressure-hydrothermal reactions were undertaken to investigate these questions.

Oxidized forms of inorganic nitrogen, thought to be the predominant non-gaseous form reaching the hadean sea, are rapidly reduced to NH3 when in contact with a variety of common metal sulfides at elevated temperature. The reaction efficiency is a function of temperature and mineral type, but reaction kinetics are extremely rapid, reaching steady state values within 1 hour.

It has always been assumed and demonstrated in the literature that organic nitrogen forms, especially amino acids, are extremely labile under high pressure, high temperature conditions. Our data indicates that while marked decreases in stability are found with increased temperature (>150°C) at any pressure, the deleterious effects of pressure upon amino acids may be in fact an artifact of the gold containers in which such studies are conducted. If such containers and their internal solutions are buffered with metal sulfides, no decrease in stability with pressure is noted going from 500 Atm. to 5000 Atm. Although not conducive to the long term stability of amino acids, HT-HP systems may not be the sites of nearly instantaneous destruction as presently thought.

What is Life?

Carol E. Cleland
Philosophy
University of Colorado
Boulder, CO
USA
Cleland@spot.colordo.edu

One of the central issues facing astrobiologists is the question "what is life?" Some researchers approach this question by trying to define "life". Yet all the popular definitions face counter-examples. This is not an accident. It reflects a fundamental confusion between two types of identity statements, definitions (e.g., bachelors are unmarried human males) and theoretical identities (e.g., water is H_20). This confusion is revealed in the nature of the definitions that have been proposed for the term "life". They invariably cite high-level properties that are used to recognize terrestrial life (e.g., complex hierarchical structure, self-regulation). This is analogous to defining water as a cooling, tasteless, odorless, wet liquid that quenches thirst. The later does not, however, provide a satisfactory answer to the question "what is water?"; it merely describes water in terms of the properties by which it is recognized by humans. Similarly, definitions of life don't provide satisfactory answers to the question "what is life?" Considerations such as this suggest that attempts to answer the question "what is life?" ought to focus on formulating plausible theoretical identities rather than definitions. Other cases of successful theoretical identity statements (e.g., water is H_20, heat is the motion of molecules) provide hints about how we might go about this.

Equilibrium Modeling of Hydrothermal Vent Fluid Cooling and its Application in a Lab-Scaled Reactor

Peter J. Dalla-Betta
Department of Geological Sciences
Arizona State University
Tempe, AZ
USA
Pdalla@asu.edu

Ken Voglesonger
Department of Geological Sciences
Arizona State University
Tempe, AZ
USA

Peggy O'Day
Department of Geological Sciences
Arizona State University
Tempe, AZ
USA

John R. Holloway
Department of Geological Sciences and Department of Chemistry and Biochemistry
Arizona State University
Tempe, AZ
USA

A catalytic reactor system was developed to operate at pressures and temperature ranges encountered at deep-sea hydrothermal vents. This system will be used to test the hypothesis that during the cooling of hydrothermal fluid, by conduction or mixing, simple organic compounds are formed metastably via catalysis by transition-metal sulfides. Hydothermal fluid will be cooled as it passes through inert (quartz) or sulfide mineral catalysts (pyrite, chalcopyrite, sphalerite, pyrrhotite). To test the affects of precipitation and dissolution of the catalyst, sulfide mineral will either be synthesized in situ from fluid cooling or placed in the reactor beforehand. In our thermodynamic model, a simplified composition (550mM NaCl, 1mM Fe+2, 10 mM HS-, 5 mM HCO_3-, pH 4.3, initial fO_2 = PPM) was used to mimic a 350 ^oC vent fluid end-member solution as it cools. A 250 bar database was created for use in the equilibrium modeling program EQ3/6. This database was composed of C_4 or smaller CH/O compounds (n-alkanes, alcohols, aldehydes, ketones, carboxylic and hydroxy acids). Additionally, pyruvic acid and several TCA cycle intermediates were also included in the model system along with CHS/O compounds (COS, CS_2, n-thiols, sulfides, and disulfides). With the formation of methane, ethane, and methanethiol inhibited, this model demonstrated to favor the formation of pyruvic acid as the major product (log activity ~ -3.2 molal), among CHO compounds, in the temperature range of 75 to 150 ^oC. Fumaric and acetic acid activities were also highest in this temperature range but generally an order of magnitude lower than that of pyruvic acid. CHO/S compounds had peak activities (log activity ethanethiol

~ -4 molal) at slightly higher temperature ranges (Tmax ~ 190°C) than the organic acids. The high theoretical activities of pyruvic, fumaric, and acetic acids indicate that some of the key components to intermediary metabolism may be formed metastably during hydrothermal fluid cooling. High activities of organic acids, in conjunction with thiols, indicate a mechanism in which key components that are needed to form thioesters may be produced. The production of thioesters is of particular importance since they have been hypothesized to be intermediates in the formation of higher-molecular weight components crucial to key metabolic pathways.

BASIC: A New Method for the Isolation of RNA Catalysts

Ico de Zwart
Department of Molecular, Cellular and Developmental Biology
University of Colorado
Boulder, CO, 80309
USA
Icenius.dezwart@colorado.edu

Mike Yarus
Department of Molecular, Cellular and Developmental Biology
University of Colorado
USA

We propose a new screening method for the isolation of catalytic RNAs from a randomized RNA pool without the need for selection or the need for catalyst self-modification. In order to find catalytic RNAs, catalytic reaction products must be related to the RNA that produced them. This will be possible if both the product and the catalytic RNA can be spatially localized in the same vicinity.

BASIC (Bead Attached Substrates for the Identification of Catalysts) is a method that meets these requirements: a substrate that yields a fluorescent product and the DNA template used to produce RNA are confined to a bead. A fluorescent bead indicates that the desired reaction has taken place, and in addition, the fluorescent bead contains template information for the catalytic RNA.

Crystallization of a Ni and FeS Protein that Fixes CO_2 into Cell Carbon

Tzanko I. Doukov
Department of Chemistry
Massachusetts Institute of Technology
77 Massachusetts Avenue, 56-546, Cambridge, MA 02139
U.S.A.
tzanko@mit.edu

Stephen W. Ragsdale
Department of Biochemistry
The Beadle Center, University of Nebraska – Lincoln
Lincoln, NE 68588-0664
U.S.A.

Catherine L. Drennan
Department of Chemistry
Massachusetts Institute of Technology
77 Massachusetts Avenue, 16-573, Cambridge, MA 02139
U.S.A.

Clostridium thermoaceticum is an anaerobic thermoacetogenic bacteria whose reductive acetyl-CoA pathway for CO_2 fixation is a remnant from the very early stages of life. In a pivotal work Wächtershäuser and Huber [1] found that a NiS and FeS converts CH_3-SH and CO into acetic acid. In the acetyl-CoA pathway methyltetrahydrofolate:corrinoid/iron-sulfur protein methyltransferase (MeTr) transfers the N^5-methyl group of (6S)-methyltetrahydrofolate to the cobalt(I) of a corrinoid/iron sulfur protein (CP). The methyl group then combines with a CO and coenzyme A to form acetyl-CoA at a Ni-FeS center of the acetyl-CoA synthase activity of the bifunctional carbon monoxide dehydrogenase/ acetyl-CoA synthase (ctCODH/ACS) enzyme [2,3].

MeTr from *Clostridium thermoaceticum* has been crystallized [4], and its structure determined by X-ray crystallography [5]. In our laboratory we are trying to obtain diffracting crystals of the other two enzymes of the pathway – CP and CODH/ACS. The latest results from the crystallization experiments will be reported.

Literature:

[1] Wächtershäuser, G. and Huber, C. (1997) Science 276, 245-247.
[2] Ragsdale, S.W. (1991) Crit Rev Biochem Mol Biol 26, 261-300.
[3] Ragsdale, S.W. and Kumar, M. (1996) Chem. Rev. 96, 2515-2539.
[4] Doukov, T., Zhao, S., Roberts, D.L., Kim, J.-J., Ragsdale, S.W. and Stezowski, J. (1995) Acta Crystallographica D51: Part 6, 1092-1093.
[5] Doukov, T., Seravalli, J, Stezowski, JJ, and Ragsdale, SW. (2000) Structure 8, 817-830.

Higher Order Models for Detecting Functional Divergence When Analyzing the Evolutionary Past

Eric A. Gaucher
Department of Molecular Cell Biology
University of Florida
PO Box 117200
Gainesville, FL 32611-7200
USA
gaucher@ufl.edu

Michael M. Miyamoto
Department of Zoology
University of Florida
USA

Steven A. Benner
Departments of Chemistry and Molecular Cell Biology
University of Florida
USA

The divergent evolution of protein sequences from genomic databases can be analyzed using different mathematical models. The most common treat all sites in a protein sequence as equally variable. More sophisticated models acknowledge the fact that purifying selection generally tolerates variable amounts of amino acid replacement at different positions in a protein sequence. In their "stationary" versions, such models assume that the replacement rate at individual positions remains constant throughout evolutionary history. "Non-stationary" covarion versions, however, allow the replacement rate at a position to vary in different branches of the evolutionary tree. Recently, statistical methods have been developed that highlight this type of variation in replacement rates. Here, we show how positions that have variable rates of divergence in different regions of a tree ("covarion behavior"), coupled with analyses of experimental three-dimensional structures, can provide experimentally testable hypotheses that relate individual amino acid residues to specific functional differences in those branches. We illustrate this in the elongation factor family of proteins using Bayesian inference. In addition, based on previous work in this laboratory (Jermann *et al., Nature* 1995), we show that incorporating higher order models of sequence evolution lead to predictions of ancestral reconstruction character states that are different than those predicted by the less-sophisticated parsimony method. Thus, we advocate the use of higher order models in comparative evolutionary analyses especially as the community attempts to predict protein function from the large amount of genomic data currently being produced by sequencing projects.

Bacterial EF-Tu and eukaryotic EF-1α are orthologous proteins involved in translation. These proteins present aminoacyl-tRNA to the ribosome for protein synthesis and are considered highly conserved based on sequence comparisons. We have analyzed these

sequences within an evolutionary context in an attempt to understand the molecular adaptations responsible for some of the known and putative functional differences between the bacterial and eukaryotic lineages. Using Bayesian statistical inferences, we were able to highlight sites that are most likely responsible for some of this functional divergence (see figure).

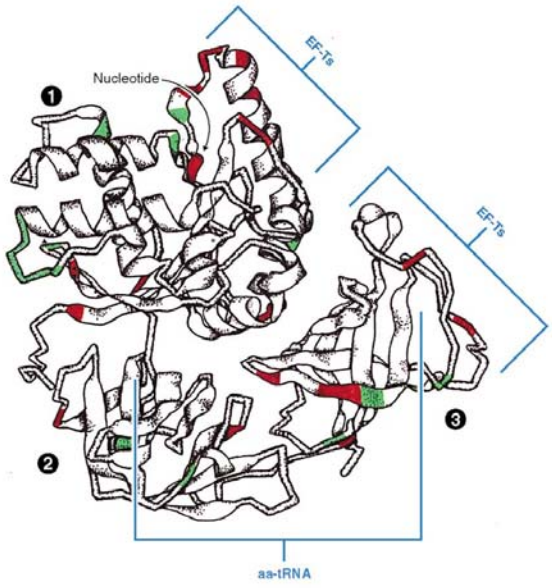


Figure legend. Tertiary structure of elongation factor. Here, green and red highlight those sites that are evolving significantly faster in bacteria than eukaryotes, and vice versa, respectively. The binding domains for the aminoacyl-tRNA, nucleotides GTP/GDP, and the nucleotide exchange factor are indicated. The shifts in functional constraints for individual sites between the two lineages signify non-stationary covarion behavior.

The Genetics of Extraterrestrial Life?

C. Ronald Geyer
Thomas R. Battersby
Steven A. Benner
Departments of Chemistry and Anatomy and Cell Biology
University of Florida, Gainesville FL 32611

One of the central questions of Astrobiology as it relates to the ongoing work of NASA to search for life through missions to other planets and their moons asks: "How would we recognize life if we were to find it?" The Watson-Crick model for DNA provides, through its elegant simplicity, the illusion of universality, especially to terrean biochemists and molecular biologists who know of no other molecule that can support Darwinian evolution.

Ultimately, organic chemists must be recruited to ask the question: "What might life, having a genesis independent from life on Earth, look like?" For more than a decade now, the Benner laboratory has asked this question, and developed a range of new organic molecules that explore possible alternative biopolymeric structures, both genetic and catalytic. From this has emerged a "second generation" model for nucleic acid structures, which places chemical constraints on the universal genetic molecule.

This talk will focus on an artificially expanded genetic information system (AEGIS), a DNA-like biomolecular system that has 12 "letters" instead of the four found in naturally occurring DNA. We have developed general rules, using the language of organic chemistry, that guides and predicts the behavior of non-standard genetic systems. These rules were developed by studying genetic behavior in a large range of DNA analogs. The results from these studies show that hydrogen bonding must be considered in any "second generation" model to predict, explain, and manipulate the behavior of nucleic acid-like artificial genetics systems. Further, we provide guidelines for designing artificial genetic systems.

Switzer, C. Y., Moroney, S. E., Benner, S. A. Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **111**, 8322-8323 (1989)
Piccirilli, J. A., Krauch, T., Moroney, S. E., Benner, S. A. Extending the genetic alphabet. Enzymatic incorporation of a new base pair into DNA and RNA. *Nature* **343**, 33-37 (1990)
Bain, J. D., Chamberlin, A. R., Switzer, C. Y., Benner, S. A. Ribosome-mediated incorporation of non-standard amino acids into a peptide through expansion of the genetic code. *Nature* **356**, 537-539 (1992)
Benner, S. A. *et al.* Redesigning nucleic acids. *Pure Appl. Chem.* **70**, 263-266 (1998)
Battersby, T. R., Ang, D. N., Burgstaller, P., Jurczyk, S., Bowser, M. T., Buchanan, D. D., Kennedy, R. T., Benner, S. A. *In vitro* selection of an adenosine receptor from a library incorporating a cationic nucleotide analog. *J. Am. Chem. Soc.* **121**, 9781-9789 (1999)
Benner, S. A. Artificial genetics. *Science* **290**, 1506 (2000)

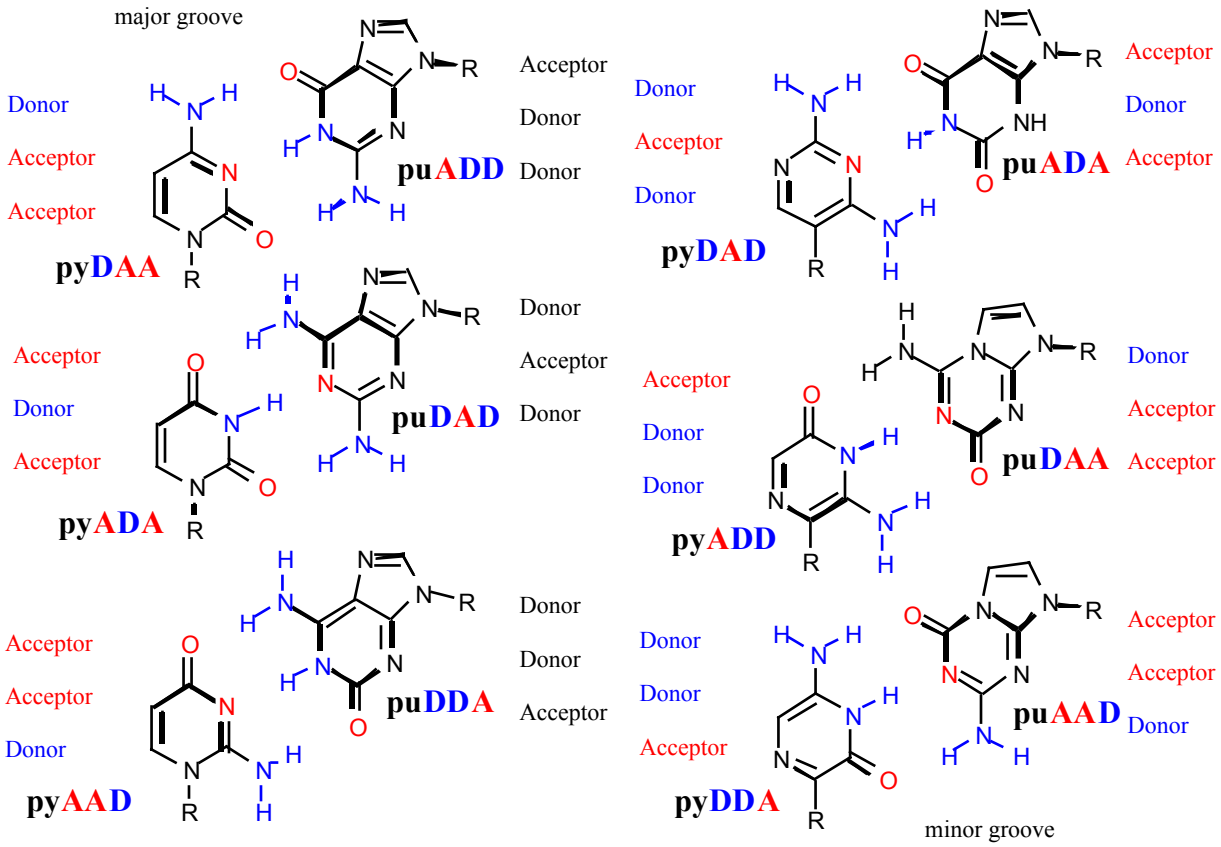


Figure. An artificially expanded genetic information system (AEGIS), a DNA-like biomolecular system that has 12 "letters" instead of the four found in naturally occurring DNA

Emergence and the Origin of Life

Robert M. Hazen
Geophysical Laboratory and NASA Astrobiology Institute
Carnegie Institution of Washington
5251 Broad Branch Road NW
Washington, DC 20015 USA
hazen@gl.ciw.edu

The geochemical origin of life may be modeled as a sequence of "emergent" events, each of which adds to molecular complexity and order. Each of these steps, if properly formulated, should be amenable to experimental study. Each emergent step, furthermore, may result in characteristic isotopic, molecular, and structural "fossils" that might be measured in extraterrestrial environments that have not been subjected to reworking by biological activity.

Natural systems with many interacting components, such as atoms, molecules, cells or stars, often display complex, "emergent" behavior not associated with their individual components. In some instances, as in the emergence of turbulent flow in fluids, the solid-state properties of crystals, or the periodic spacing of sand dunes, such complex behavior can be modeled *a posteriori* when appropriate interaction parameters have been determined. Other phenomena, such as the emergence of consciousness from collections of neurons or the emergence of social behavior from collections of humans, are, at least for the present, less amenable to quantitative analysis.

This observed emergent behavior of highly ordered systems, including galaxies, planets, and life, points to a universal organizing principal. Natural systems tend to develop local-scale order – non-equilibrium regions of spontaneously increased free energy and decreased entropy – even as global-scale entropy increases. A familiar non-biological example is provided by the formation of sand dunes. Gradual deposition of sediments in a lake or deltaic environment may lead to flat-flying sand deposits with poorly-sorted grain sizes. Such a sediment system has minimum gravitational potential energy and maximum configurational entropy. If, however, that sediment system dries and is subjected to a steady wind, new structures will emerge. Sand dunes form with a variety of periodic shapes, while sediment grains become highly sorted by size. As wind energy flows across the sediment-atmosphere interface, the sediments spontaneously gain gravitational potential energy, while their configurational entropy decreases.

This universal tendency for systems to display increased order at an energy-rich interface, while consistent with the first and second laws of thermodynamics, is not formally addressed in either of those laws; indeed, some researchers have proposed that this behavior should be systematized in a "fourth law of thermodynamics."

If the chemical evolution of life occurred as a sequence of successively more complex stages of emergence, then the divide between non-life and life may ill-defined. One might establish a hierarchy of emergent properties – a progressive sequence that leads, for example, through a number of steps from a pre-biotic ocean enriched in organic

molecules, to a cluster of molecules arranged on a mineral surface, to self-replicating molecular systems, to encapsulation and eventually prokaryotic life. The exact nature and sequence of these steps may vary in different environments, but any definition that distinguishes between non-living and living systems of necessity becomes more arbitrary as the number of discrete emergent steps to life increases.

The concept of a sequence of discrete emergent steps is useful and appealing in experimental and theoretical studies of the origin of life for at least two pragmatic reasons. First, a progression of steps reduces an immensely complex historical process to a succession of several, more manageable chemical episodes. Each step becomes a focused process for laboratory experimentation or theoretical modeling. Second, each of these steps may result in distinctive, measurable isotopic, molecular, and structural signatures. As we search for life elsewhere in the universe, we may thus be able to characterize extraterrestrial environments according to their degree of emergence along this multi-step path. In this context, it is useful to review experimental programs that attempt to elucidate a few of the possible geochemical steps in the emergence of life.

Hydrothermal Reactions of Pyruvate: Production of Amphiphilic Molecules and Vesicle Formation

Robert M. Hazen[1]
George D. Cody
Hatten S. Yoder, Jr.
Geophysical Laboratory and NASA Astrobiology Institute
Carnegie Institution of Washington
5251 Broad Branch Road NW
Washington, DC 20015 USA
[1]hazen@gl.ciw.edu

David W. Deamer
Department of Chemistry and Biochemistry
University of California
Santa Cruz, CA 95064 USA

The synthesis of amphiphilic molecules capable of aqueous self-assembly into membrane-like structures, including bilayers and micelles, is an essential step in the emergence of a protocell. Mechanisms for the prebiotic synthesis and assembly of amphiphiles are thus of considerable interest.

We observe amphiphile synthesis from pyruvate in a CO_2-water fluid subjected to temperatures from 250°C to 350°C and pressures from 0.05 to 0.2 GPa (500 to 2000 atmospheres) for 2 hours. Principal run products include acetic acid and CO_2 (from decarbonation of pyruvate) and methyl succinate (from dimerization of pyruvate and subsequent decarboxylation). We also observe up to 50% conversion of pyruvate to a water-insoluble, yellow-brown, strongly aromatic oily residue. Analysis by GCMS reveals this material to be a complex suite, including cyclic aromatic compounds, presumably formed by polymerization and subsequent cycloaddition reactions.

We characterized this material by extracting and analyzing the chloroform-soluble fraction. We examined this material by 2D TLC chromatography, which revealed a pattern of seven distinct fluorescent regions – a pattern that is qualitatively similar to that observed for amphiphilic components of the Murchison meteorite by Deamer and Pashley ("Amphiphilic components of the Murchison carbonaceous chondrite: Surface properties and membrane formation." *Origins of Life and Evolution of the Biosphere* **19**, 21-38, 1989). Each of these seven TLC areas was extracted, dried, washed in a phosphate buffer (pH = 8.5), and examined by fluorescence microscopy. One of these regions consisted of a significant fraction of surface-active molecules, which form apparent monomolecular films at air-water interfaces. When placed in the phosphate-buffered aqueous solution, these molecules also organize into fluorescent, membranous vesicle-like structures. We conclude that hydrothermal processes, perhaps similar to those that occurred on the Murchison parent body, may lead to the production of amphiphilic, membrane-forming molecules.

Functional Proteins from a Random-Sequence Library

Anthony D. Keefe and Jack W. Szostak
Harvard Medical School Department of Genetics, and
Massachusetts General Hospital Department of Molecular Biology, and
Howard Hughes Medical Institute
Boston, MA 02114 USA
keefe@molbio.mgh.harvard.edu

Functional primordial proteins presumably originated from random sequences, but it is not known how frequently functional, or even folded, proteins occur in collections of random sequences. To address these issues we used *in vitro* selection of mRNA-displayed proteins to allow sampling of a large number of distinct random sequences. Starting from a library of $6x10^{12}$ proteins each containing eighty contiguous random amino acids, we selected functional proteins by enriching for those that bind to ATP. This selection yielded four novel ATP-binding proteins that appear to be unrelated to each other or to anything found in the current databases of biological proteins. The frequency of occurrence of functional proteins in random-sequence libraries appears to be similar to that observed for equivalent RNA libraries.

Keefe, A. D. and Szostak, J. W. Functional Proteins from a Random-Sequence Library. *Nature* (2001); in press

Evolutionary Accretion of Small Motifs Modulates RNA Activity

Rosa Larralde
Department of Molecular Biology
Massachusetts General Hospital / Harvard Medical School
50 Blossom St, Wellman 9; Boston, MA 02114
USA
larralde@genetics.mgh.harvard.edu

Jack W. Szostak
Department of Molecular Biology
Massachusetts General Hospital / Harvard Medical School
50 Blossom St, Wellman 9; Boston, MA 02114
USA

During *in vitro* selection experiments, it has become apparent that the simplest (i.e. shortest) solutions, not necessarily the best performing ones, tend to dominate and take over the surviving populations. This can be understood in the light that simple motifs are usually much more frequent than complex ones in starting populations and they will remain so in succesive generations as long as they are fit enough to satisfy the selective pressure.

It is not so clear what happens to simple solutions in the context of a changing environment, scenarios in which the selective pressure, for example, sets a new, more stringent functional cut-off for survival. If a simple, dominant solution is challenged to the limit of its fitness, can it be somewhat improved upon, or would it have to be abandoned and transformed into a different and more apt design?

To address this question, the simplest ATP-aptamer known (the Sassanfar motif) was placed in a context of extra random sequence, and taken through an *in vitro* selection aiming for tighter ATP-binding. The change of selective pressure upon pre-existing functionalities has lead, in this instance, not to the establishment of a totally new solution, but to the recruitment of very small motifs that seem to modestly modulate the activity of the pre-existing aptamer.

This observation seems to indicate that very large and complex functional RNAs need not have been established all at once in answer to particular chemical needs, but they may have slowly increased in size and varied their activity through the addition of small modules.

Synthesis of Alpha-Amino Acids in Hydrothermal Media from Selected Primitive Starting Compounds

Kono H. Lemke
Department of Geological & Environmental Sciences
Stanford University
klemke@pangea.stanford.edu

David S. Ross
U.S. Geological Survey
Menlo Park

James L. Bischoff
U.S. Geological Survey
Menlo Park

Dennis K. Bird
Department of Geological & Environmental Sciences
Stanford University

The source of amino acids from primitive starting substances remains a matter of conjecture, and the commonly employed starting mix in laboratory studies of prebiotic synthesis includes formaldehyde and HCN. Thermodynamic calculations for a CO_2/N_2/liquid water system with iron oxide/silicate or iron oxide/sulfide redox buffers show, however, that suitable levels of formaldehyde and cyanide cannot be present. Our calculations show, rather, that NH_4^+, HCO_2^- and higher carboxylic acids are strongly preferred. Thus either formaldehyde and HCN were externally supplied at sufficient rates, or early amino acids were derived from NH_4^+ and HCO_2^-. Formate is an attractive candidate for the starting mixture since it is the basis of Fischer-Tropsch synthesis, and therefore carbon-carbon bond formation.

The study described here employs various combinations of formaldehyde, HCN, ammonium, and formate in liquid water at 210°C. The goal is to ascertain the minimum set of starting materials for amino acid synthesis.

Here we present results studying generation alpha-amino acids by heating of various combinations of ammonia, formaldehyde, formate and cyanide at 210˚C at vapor saturation over a period of 3 hours. The synthesis products include glycine, alanine, serine and aspartic acid. In our first study we observe that 1-2% of total carbon is converted from aqueous ammonium cyanide and formaldehyde to glycine with total alanine, serine and aspartic acid yields below 2% of total glycine. Synthesis experiments with formaldehyde alone show no measurable quantities of amino acids. Reactions without formaldehyde, but with ammonium cyanide yield glycine as the most abundant acid at 0.02% of total carbon. In the same experiment we find alanine and aspartic acid yields at 40% and at 8% of glycine production, respectively. There is suggestive evidence

that in the absence of starting aldehyde glycine synthesis occurs in a manner akin to Fischer-Tropsch synthesis, namely via the disproportionation of formate to carbon dioxide and amino acids. These results demonstrate that aldehydes are not necessary starting reactants on the route to amino acids. An alternative pathway is presented here with Fischer-Tropsch chemistry at its core where the principal carbon source formate is disproportionated over a natural catalyst.

Finally, it is important to consider the limitations to conclusions to be drawn from laboratory studies at elevated temperatures conducted over relatively brief periods, but extending perhaps to weeks or months, as applied to prebiotic activity at more modest temperatures over periods of millions of years. First, examination of a range of Arrhenius parameters anticipated for the pertinent chemistry shows that reactions with measurable rates at 200°-300°C will operate sensibly at 25°C with characteristic times in the range of 10^6 to 10^7 years. Thus workers may properly infer a link to prebiotic settings from successful observations of amino acids (or other molecules of prebiotic interest) in hydrothermal laboratory studies.

On the other hand the absence of amino acids in the laboratory cannot necessarily lead to judgements that the chemistry in question was not prebiotically operable. First, the thermodynamic favor that might be present at the more modest condition could vanish at higher temperatures. (This situation was shown to be the case for the FeS-H_2S/FeS_2 system recently by Schoonen, et al. (1999).) Second, and perhaps more significant, is the likely imposition of degradative reactions with relatively high activation energies at the more severe laboratory conditions - decarboxylation for example - which would not be competitive in the prebiotic environment. Care must be taken in carrying out the extreme extrapolations we must engage in in this line of work, and accordingly both the thermodynamic and chemical kinetic arenas must be evaluated in such studies.

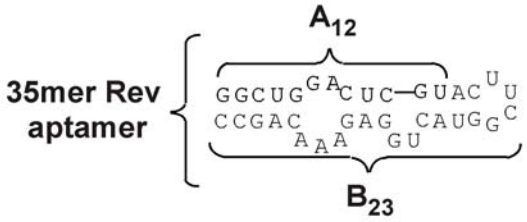
Molecular Recognition, Catalysis and the Origin of Life

Matthew Levy
Institute for Cellular and Molecular Biology
University of Texas at Austin
Austin, Texas 78712
USA
mattlevy@mail.utexas.edu

Andrew D. Ellington
Institute for Cellular and Molecular Biology
University of Texas at Austin
Austin, Texas 78712
USA

In an effort to better understand how the principles of molecular recognition may have affected the origin of biocatalysis on Earth, we have been using two different systems to investigate the ability of peptides and small molecules to facilitate the ligation of oligonucleotides.

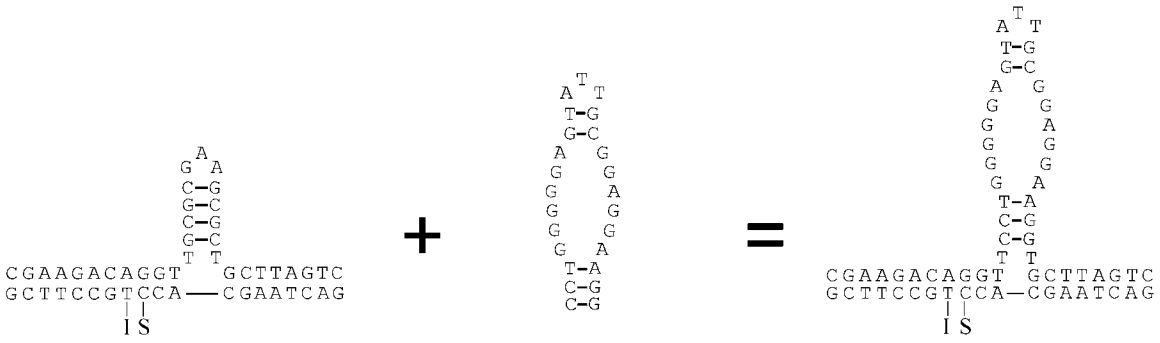
In the first of theses two systems, we have been attempting to reassemble fragmented RNA aptamers in the presence of their cognate peptide ligands. We have found using the activating agent cyanogen bromide that ligation of two RNA aptamer fragments is specifically enhanced up to ~10 fold in the presence of the aptamers cognate peptide. However the overall yields for these ligation reactions are low (~1-5%), and the reactions are plagued by the formation of side products.



17mer Rev peptide **TRQARRNRRRRWRERQR**

In an effort to increase ligation yields and avoid the complications of side reactions, we have also been investigating an alternative system based a deoxyribozyme ligase previously isolated in our laboratory which can catalyze a ligation reaction between an oligonucleotide bearing a 5'iodine and another containing a 3' phosphorothioate.

Structure analysis of this deoxyribozyme as well as deletion constructs have indicated that while much of the catalytic potential of this ribozyme was derived from templating, about 10 fold could be attributed to the presence of a 3-way junction at the ligation junction. The presence of this structural feature also suggested that the ligation reaction might be attenuated by changes in the structure of this addition stem-loop. To test this hypothesis we replaced this stem in a minimal trans-ligating deoxyribozyme with an ATP binding DNA aptamer. The resulting DNA-aptamer construct (a deoxyaptazyme) showed a 250 fold increase in ligation rate in the presence of its cognate ligand (ATP).



Minimal trans-ligation construct | DNA ATP aptamer | Effector dependent DNA template

These results further support the idea that small molecules can modulate the interactions of nucleic acids as has been suggested by previous work in our laboratory and others, with the evolution and design of effector dependent ribozymes. In addition they provide a means for the design and study of effector dependent cyclic and hypercyclic networks, a project which we are currently pursuing in our lab.

Gene Fusion: A Genome Wide Survey

Ping Liang
Josephine Bay Paul Center for Evolution and Comparative Molecular Biology
Marine Biological Institute
7 MBL Street
Woods Hole, MA 02543
USA
pliang@mbl.edu

Monica Riley
Josephine Bay Paul Center for Evolution and Comparative Molecular Biology
Marine Biological Institute
7 MBL Street
Woods Hole, MA 02543
USA

As a well known fact, organisms form larger and complex multimodular (composite or chimeric) and mostly multi-functional proteins through gene fusions of two or more individual genes which have independent evolution histories and functions. We call each of these components a module. The existence of multimodular proteins may improves the efficiency in gene regulation and in cellular functions, and thus may give the host organism advantages in adaptation to environments. Analysis of all gene fusions in present-day organisms should allow us to examine the patterns of gene fusion in context with cellular functions, to trace back the evolution processes from the ancient smaller and uni-functional proteins to the present-day larger and complex multi-functional proteins, and to estimate the minimal number of ancestor proteins that existed in the last common ancestor for all life on earth. Although many multimodular proteins have been experimentally known, identification of gene fusion events systematically at genome scale had not been possible until recently when large number of completed genome sequences have been becoming available. In addition, technical difficulties for such analysis also exist due to the complexity of this biological and evolutionary process. We report from this study a new strategy to computationally identify multimodular proteins using completed genome sequences and the results surveyed from 22 organisms with the data from over 40 organisms to be presented during the meeting.

Methods: To identify modules, a protein sequence database containing all protein sequences for each complete genome involved in the study was built, and a all-against-all sequence similarity search was performed using DARWIN (Data Analysis and Retrieval With Indexed Nucleotide/Peptide Sequences) program or NCBI blastp program. The alignment positions and the sequence similarity data were extracted for each sequence match. To focus on the recombination involving only whole individual proteins and to avoid the complicated situations due to the recombination of proteins at motif and domain levels which makes a systematic approach almost impossible, identification of a module is limited to the following situations: 1) a part of a protein or a whole protein sharing sequence similarity with another protein at full length; 2) one protein sharing sequence similarity at N-terminal region with another protein at C-terminal region or vice

versa. In addition, all modules have to be at least 83 residues in length with the sequence similarity used for identifying the module no more than 200 in PAM distance (for DARWIN) or less than 0.0001 in E value (for blastp). A further step of this analysis was to assemble all proteins into both paralog or paralog plus ortholog families based on the sequence relationship for all identified modules.

Results: 1) Distribution of multi-modular proteins among all studied genomes. Among the 22 genomes which include 14 eubacteria, 6 archaea, and 2 eukaryotes (S. cerevisiae and C. elegans), the frequency of multimodular proteins ranges from 17% to 40%. Although there is no clear relationship between the frequency of multmodular proteins and the genome or proteome size among the prokaryotes, eukaryotes being the two largest genomes on the list have higher percentage of multimodular proteins (38% for yeast and 40% for worm) than any prokaryotes. Among all bacteria, Halobacterium sp. and Borrela burgdorferi have the highest percentage of multimodular proteins (35%), while Aeropyrum pernix has the lowest (17%). 2) Relationship between gene fusion and cellular functions. The 850 multimodular proteins from E. coli K-12 were used to conduct this study since this genome has most function annotation available. Among the 9 categories of proteins, transporter proteins have the largest number, as well as the highest percentage, of proteins as multimodular proteins (47%), followed by membrane proteins (33%), and then by transcription regulators (24%). Enzymes have a relative lower percentage of multimodular proteins (18%), while structure proteins have the lowest among all known categories (7%). Whether or not this pattern is shared by other genomes awaits further studies. Demonstrations of gene fusion as a dynamic process will be presented in poster using examples from 4 closely related species, two E. coli strains (K-12 and O157:H2), Salmonella typhymurium, and Pseudomonas aeruginosa.

Conclusions: Gene fusion occurs widely among all organisms with higher frequencies in eukaryotes than in prokaryotes. No clear relationship between this gene fusion frequency and genome/proteome size were observed. Based on the data from E. coli, the frequency of gene fusion differs among different function categories, with transporter proteins having the highest and structural proteins having the least, reflecting the different nature of cellular functions. Data also suggest gene fusion is a dynamic process that differs from genome to genome and it represents the adaptation of organisms to their specific physiological requirements.

This work was supported by NASA Astrobiology Institute grant NCC2-1054.

In-Situ Spectroscopic Observation of Abiotic Chemical Reactions Producing Organic Matter

Takafumi Matsui
Department of Complexity Science and Engineering
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo
JAPAN
Matsui@eps.s.u-tokyo.ac.jp

Seiji Sugita
Department of Earth and Planetary Science
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo
JAPAN

Hiroshi Imanaka
Department of Earth and Planetary Science
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo
JAPAN

Masanori Fuyuki
Department of Complexity Science and Engineering
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo
JAPAN

Kensei Kobayashi
Department of Chemistry and Biotechnology
Yokohama National University
Hodogaya-ku, Yokohama
JAPAN

One of the most important issues in the origin of life is how organic matter was supplied to the surface of the Earth. Proposed mechanisms so far can be categorized generally to two types: endogenous and exogenous. The former is organic synthesis from inorganic components of either the atmosphere or the ocean with high-energy processes, such as lightening, impact-induced shock waves, and cosmic ray radiation. The latter is delivery of organic matter from the space through infall of asteroids, comets, meteors, and other bodies to the Earth. Although most attention is paid to the surviving organic matter through the exogenous mechanisms, most of organic matter reached the atmosphere or the surface of the Earth is expected to be destroyed upon its arrival because of the extremely high temperature (i.e., high thermal energy). Consequently, no matter which types of mechanism is dominant, high-energy chemical reactions play a crucial role in supply of organic matter on the prebiotic Earth. However, little direct observation has been made for such high-energy chemical reactions in laboratory experiments in the context of the origin of life. Since the physical conditions for high-energy chemical reaction is usually very dangerous and transient, most measurement methods cannot be

used. Thus, the high-energy chemical reactions have been treated as “black boxes”. Nevertheless, optical spectroscopy may provide a “peep hole” to the black boxes. Because of the high level of energy, these chemical reactions are often associated with intense light emission. The spectrum of the light emission reflects the physical and chemical conditions of the reaction field, and it has practically no time delay between change in the physical/chemical conditions and that in emission spectra.

The purpose of our study is to understand the high-energy chemical reactions through in-situ spectroscopic observations. We have been observing many different processes, such as electric discharge (i.e., lightening) using both Tesla coils and pulse lasers, shower of high-energy charged particles (i.e., cosmic ray radiation) using a van de Graaf accelerator, and hypervelocity impacts (i.e., collisions of asteroids and comets) using both two-stage light gas guns and an electromagnetic rail gun. Each process shows strong and characteristic emission spectrum. Preliminary analysis indicates that these spectra provide the temperature, density, chemical composition, and departure from a thermal equilibrium of the chemical reaction field. For example, both impact-induced vapor and simulated lightening show emission spectra consistent with a local thermal equilibrium, whereas simulated cosmic ray radiation shows a large departure from a thermal equilibrium. Such difference may account for the fact that radiation of charged particles can produce organic matter very efficiently even in an oxidizing atmosphere (e.g., CO_2-N_2-H_2O system).

Our experiments are still in a preliminary phase. Most of the emission spectra taken so far do not have precise time resolution or spatial resolution. We are currently setting up a high-speed spectroscopic measurement system, which can capture phenomena with a time scale of 10 ns and has one-dimensional spatial resolution. Both temporal and spatial variation of emission spectra observed by the new system will permit quantitative characterization of the high-energy chemical reactions and will reveal their significance in the origin of life.

Multiple Catalytic Functions of the *Escherichia coli* Fad B Protein Terminal Region: Anatomy of a Promiscuous Active Site

Tom McCormack
7 MBL Street, Woods Hole, MA
The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution
Marine Biological Laboratory
USA
tmccormack@mbl.edu

Monica Riley
7 MBL Street, Woods Hole, MA
The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution
Marine Biological Laboratory
USA

In *Escherichia coli*, the fadBA operon encodes the fatty acid oxidation multienzyme complex. While the C-terminus of the FadB protein carries out a single reaction, NADH-dependent dehydrogenation, the N-terminal region, which shares homology with a wide range of CoA-binding enzymes, performs three separate catalytic functions, namely isomerization, hydration and epimerization. At least two of these reactions, epimerization and hydration, appear to occur at a single active site defined by two catalytic glutamate residues, while the third reaction, isomerization appears to possess an active site with elements both common to and separate from the active site of the other two reactions.

The three-dimensional structures of several proteins from the hydratase/isomerase superfamily have been solved, providing much information about the active site residues and catalytic mechanisms of enzymes within a superfamily which shows considerable variability in both substrate specificity and reaction chemistry. Structural alignments reveal a strikingly similar active site fold in spite of a diverse combination of catalytic residues for different enzymes within the hydratase/isomerase superfamily. Among the proteins within this superfamily whose structures have been determined, the *E. coli* FadB N-terminal domain shares greatest sequence similarity to rat liver enoyl CoA hydratase, allowing us to construct a three-dimensional model of the FadB N-terminus. FadB illustrates the diversity of mechanisms by which proteins can perform multiple functions, even within a common active site motif.

MALDI-MS, HPLC-APCI-MS and Solids NMR Analysis of Hydrogen Cyanide Polymers

R.D. Minard
Penn State Astrobiology Research Center and Department of Chemistry
Penn State University
University Park, PA, 16802.
UNITED STATES
rminard@psu.edu

M.N. Fiddler, E.P. Shields, K.T. Mueller, A.D. Jones
Department of Chemistry
Penn State University
University Park, PA, 16802.
UNITED STATES

A.L. Eastwood
Department of Chemistry
University of Virginia, Charlottesville, VA
UNITED STATES

J. Schendel
Department of Chemistry
University of Puget Sound, Tacoma, WA
UNITED STATES

The role of hydrogen cyanide polymer chemistry in the origin of life has provoked much speculation. In spite of extensive efforts by many groups, a satisfactory understanding of this polymer's structure and mechanism of formation still eludes us. Our studies involve the application of modern powerful instrumental methods to this problem. Using MALDI-MS we have found that hydrogen cyanide polymers, $(HCN)_n$, formed from HCN monomer, showed MH^+ peaks for n = 8 to 25 with the maximum occurring at n = 15 to 17. HCN polymer formed from diaminomaleonitrile (HCN tetramer) showed a similar MALDI spectra with no enhancement for peaks at n = 4, 8, 12, 16, etc. implying the breakdown of the tetramer to monomer followed by monomer repolymerization. The MALDI-MS spectra of polymer formed from aminomalononitrile (HCN trimer) showed oligomers from m/z 200 to 800, but these did not match an $(HCN)_n$ series.

In other studies, the water soluble products derived from stirring HCN polymer with water at room temperature for 1 to 3 days consisted primarily of two compounds, urea and a compound with formula $C_6H_8N_6$ as determined by APCI-exact mass MS whose structure we are determining. We are also carrying out the synthesis of $(H^{13}C^{15}N)_n$. Solids $^{13}C/^{15}N$ NMR analysis of this labeled polymer should provide considerable structural detail of HCN polymer.

Molecular Dynamics (MD) Lattice Gas for Modeling Molecular Self-Assembly and Self-Organization Processes

Martin Nilsson and Steen Rasmussen
Los Alamos National Laboratory
Los Alamos NM 87545
USA
[nilsson, steen @lanl.gov]

This new simulation method fills a gap between traditional molecular dynamics (MD) simulations on the one hand and lattice gas automata (LGA) and lattice Boltzmann equations (LBE) on the other. MD simulations generally address molecular processes studied on length scales of angstroms to nanometers and over times scales of pico- to nanoseconds. LGA and LBE simulate complex fluids and fluid flow phenomena on length scales of submicrometers to meters and over time scales of microseconds to minutes. Our MD lattice gas simulations address molecular interactions on nanometer to micrometer length scales and over time scales up to milliseconds on workstations and seconds on supercomputers. This intermediate range suits them to modeling molecular self-organization and self-assembly processes involving ions, monomers, complex polymers, polymer aggregates (supermolecular structures), complex surfaces with charge, and chemical reactions.

With our MD lattice gas, we have simulated molecular interactions spanning three levels of complexity. At the first level, we begin with water molecules, ions, and both hydrophilic and hydrophobic monomers as our basic ingredients. At the second level in a aqueous environment, monomers can form amphiphilic chains (i.e., polymers with a hydrophilic head and hydrophobic tail). At the third level, these polymers can self-assemble into amphiphilic aggregates such as micelles or vesicle membranes, which under certain conditions can self-reproduce. At each level, we observe how local interactions generate higher-order structures with new functionalities not exhibited at the level of their constituents.

As the name indicates, our MD lattice gas is defined in a discrete space. Calculations involve information particles that move on two superimposed 3-D lattices, a molecular lattice and a field lattice. On the molecular lattice, particles carry explicit information about the structure of matter—for example, whether it consists of a water molecule, a hydrophobic or hydrophilic monomer, or a reactive radical. Particles (molecules) have excluded volumes: only one can reside at a lattice site at any given time. They can also have an orientation: as neighboring molecules interact, they rotate to find a local minimum in terms of their potential energy. Finally, molecules have an associated kinetic energy in each lattice direction that changes as they collide with one another.

On the field lattice, particles carry explicit information about the molecules' electromagnetic field structure, which determines how the molecules interact. For example: A water molecule has four well-defined hydrogen bonds that influence its

movement and dictate how it interacts with other molecules. Our dynamics is driven by six intermolecular interactions: dipole–dipole, charge–charge, hydrogen bond, dipole–induced dipole, induced dipole–induced dipole, and cooperativity. Each molecular interaction is thus decomposed into a set of repelling and attracting particles of varying values depending on the type and position of the interaction. These interactions (potential energy terms) account for the physicochemical properties of our molecular species that are crucial to the species' self-assembly in a polar environment.

The MD lattice gas interactions conserve mass, energy, and momentum and are defined by the truncated Schrödinger equation. The resulting field at any lattice site influences the resident molecule and determines where it moves next on the molecular lattice. If no molecule is present at a lattice site, no force fields is propagated from that site. However, fields can reach and pass through unoccupied sites as well as be influenced (partially shielded) by the molecules at occupied sites.

Covalent bonds between monomers in polymers are also viewed as information particles that must be propagated so that the bonds do not break as the polymer moves around on the lattice. Only local interactions for molecules are used to move extended objects such as polymers and aggregates. That is, our model follows a bottom-up approach: interactions are derived from the laws of physics, which have been modified only to accommodate the constraints of a 3-D lattice. A more-detailed explanation of our MD lattice gas is given in the technical paper in the Appendix.

The molecular dynamics in our simulations are determined by three sets of rules: (1) rules that propagate information particles on the field lattice, (2) rules that evaluate this received information together with the local molecular state (i.e., the molecular type, orientation, and kinetic energy), and (3) rules that move particles on the molecular lattice and transform the system into the next time step. Since our system operates on superimposed lattices and contains several different types of information particles, the update cycle is more complicated than that of traditional lattice gases.

In LGA simulations, the update cycle consists of two steps: molecular collisions followed by molecular motion. One way to view this two-step cycle is as "tick-tock," where *tick* refers to a collision and *tock* refers to the molecules' resulting moves. By contrast, our simulations incorporate more detailed interactions—e.g., force-field propagation, rotations, and chemical reactions are tracked in addition to collisions. Our update cycle therefore goes "tick-tick- . . . tick-tock." A sequence of "ticks" primarily on the field lattice result in a "tock" on the molecular lattice. Splitting the update cycle into a series of physical processes (with differing time scales) that eventually move the molecules results in computational simplicity (modularity and suitability for parallel processing) and means that each step in our molecular dynamics has a clear physical interpretation.

In short, the microscopic interactions we track are based on first principles and are both deterministic and reversible. The overall simulation corresponds to a microcanonical ensemble. We have used our MD lattice gas to model micelle formation, the hydrophobic effect, phase separations, hydrocarbons in water, amphiphilic fluids, complex fluids at mineral interfaces, self-reproducing micelles, membrane stability, and the dynamics of templating polymers such as RNA and PNA (ribonucleic and peptide nucleic acids).

Models of Protocellular Structure, Function and Evolution

Michael H. New
Exobiology Branch
NASA Ames Research Center
Moffett Field, CA 95035-1000
USA
mnew@mail.arc.nasa.gov

Andrew Pohorille
NASA Center for Computational Astrobiology
NASA Ames Research Center
USA

Jack W. Szostak
Department of Molecular Biology
Massachusetts General Hospital
USA

Tony Keefe
Department of Molecular Biology
Massachusetts General Hospital
USA

Janos K. Lanyi
Department of Physiology and Biophysics
University of California, Irvine
USA

In the absence of any record of protocells, the most direct way to test our understanding of the origin of cellular life is to construct laboratory models that capture important features of protocellular systems. Such efforts are currently underway in a collaborative project between NASA-Ames, Harvard Medical School and University of California. They are accompanied by computational studies aimed at explaining self-organization of simple molecules into ordered structures.

The centerpiece of this project is a method for the *in vitro* evolution of protein enzymes toward arbitrary catalytic targets. A similar approach has already been developed for nucleic acids in which a small number of functional molecules are selected from a large, random population of candidates. The selected molecules are next vastly multiplied using the polymerase chain reaction. A mutagenic approach, in which the sequences of selected molecules are randomly altered, can yield further improvements in performance or alterations of specificities. Unfortunately, the catalytic potential of nucleic acids is rather limited. Proteins are more catalytically capable but cannot be directly amplified. In the new technique, this problem is circumvented by covalently linking each protein of the initial, diverse, pool to the RNA sequence that codes for it. Then, selection is performed on the proteins, but the nucleic acids are replicated.

To date, we have obtained "a proof of concept" by evolving simple, novel proteins capable of selectively binding adenosine tri-phosphate (ATP). Our next goal is to create an enzyme that can phosphorylate amino acids and another to catalyze the formation of peptide bonds in the absence of nucleic acid templates. This latter reaction does not take place in contemporary cells.

Once developed, these enzymes will be encapsulated in liposomes so that they will function in a simulated cellular environment. To provide a continuous energy supply, usually needed to activate the substrates, an energy transduction complex that generates ATP from adenosine diphosphate, inorganic phosphate and light will be used. This system, consisting of two modern proteins, ATP synthase and bacteriorhodopsin, has already been built and shown to work efficiently. This system has also been coupled to a chemical reaction: the production of acetyl-coenzyme A from acetate and coenzyme A. The reaction was catalyzed by a thermophilic acetyl-CoA synthetase that utilizes ATP, but not ADP, to make acetyl-CoA. After adding this enzyme and appropriate substrates to liposomes containing the ATP synthase and bacteriorhodopsin, net synthesis of acetyl-CoA was observed upon illumination.

We are further actively working towards simplifying the bio- energetics systems. At present, our efforts are concentrated on designing a simple, reversible, photo-induced proton pump using extensive, atomic-level computer simulations. Three different mechanisms of proton pumping have been proposed. One system that is being examined in detail is a proton channel formed by the transmembrane portion of the M_2 protein of influenza virus. This system appears to be an excellent candidate for re-engineering into a proton pump.

Protocells must have had an ability not only to self-maintain but also to evolve. Previous studies on protocellular evolution have focused on self-replication. In these systems, Darwinian evolution occurs through a series of small alterations to functional molecules whose identities are stored. Protocells, however, may have been incapable of such storage. We hypothesize that under such conditions, the replication of functions and their interrelationships, rather than the precise identities of the functional molecules, is sufficient for survival and evolution. This process is called non-genomic evolution.

Recent breakthroughs in experimental protein chemistry have opened the gates for experimental tests of non-genomic evolution. On the basis of these achievements, we have developed a stochastic model for examining the evolutionary potential of non-genomic systems. We have demonstrated that the catalytic capabilities of such a system can increase with time, which means that the system can evolve. Its evolutionary potential is dependent upon the competition between the formation of bond-forming and bond-cutting catalysts. We will further discuss how the system can develop new functions and how these functions can couple to increase metabolic capabilities of protocells.

Fate of CO Hydration Products in Anoxic Solution: The Photochemical and Thermal Decomposition of Na-Formate

Robin Penfield
Department of Geosciences
State University of New York - Stony Brook
Stony Brook, NY 11794-2100
USA
rpenfiel@ic.sunysb.edu

Martin A. A. Schoonen
Department of Geosciences
State University of New York - Stony Brook
Stony Brook, NY 11794-2100
USA
mschoonen@notes.cc.sunysb.edu

James F. Kasting
Department of Geosciences
Pennsylvania State University
211 Deike Building, University Park, PA 16802
USA
kasting@essc.psu.edu

Robert D. Minard
Department of Chemistry
Pennsylvania State University
152 Davey Laboratory, University Park, PA 16802
USA
rminard@psu.edu

The thermal and photochemical decomposition of sodium formate has been measured for dilute aqueous solutions over a range of temperatures (4^o – 75^oC), pH (5.4 - 9.0), and radiation sources (253.7 nm, 800-2500 nm, and darkness). Formaldehyde abundance was determined through spectrophotometry, and both sodium formate and acetate were measured through ion chromatography. This rate data will be used to evaluate carbon cycling between the early atmosphere and hydrosphere, and the potential accumulation of simple organics such as formate and formaldehyde in an early ocean. The secondary reactions of formate in the ocean may have been relevant to prebiotic synthesis. For instance, sodium formate disproportionates to formaldehyde and sodium bicarbonate in the absence of oxygen (Sinnarkar and Ray, 1975; Elliot et al. 1983), and could potentially lead to formose polymerization if formaldehyde concentrations approached 10mmol/L (Mayer et al.1963).

In anoxygenic, ambient pressure enclosures, we find that the kinetics of formate decomposition is first order in formate. UV radiation has a significant effect on the rate constant **k**: log**k** = -7.30 +/- 0.62 s^{-1} without light at 35^oC, but increases to -5.38 +/- 0.12

s^{-1} with a 800-2500 nm Xe lamp, and to 3.62 +/- 0.081 s^{-1} with a 253.7 nm Hg lamp. Formaldehyde is a transient species in the far-UV experiments, and may serve as an intermediate in the production of acetate. Low levels of acetate appear in solution, at rates that increase with solution pH (0.42% of formate carbon is converted to acetate at pH 6.5 vs. 3.5% at pH 9). For non-irradiated solutions, the Arrhenius relation shows an E_A of 83.80 +/- 4.68 kJ/mol and pre-exponential factor of 15.6 +/-1.78. Preliminary data for 4°C and 50°C suggests a 1:1 ratio of formate consumption to formaldehyde production in the non-irradiated solutions. If this relationship is consistent at all temperatures, we would propose that formate decomposition proceeds via a first step disproportionation to formaldehyde and bicarbonate in the absence of oxygen. The acetate that forms in basic solutions in the presence of intense UV radiation may be the result of a Cannizzaro-type reaction (2) to methanol and formate (Matsuura and Smith, 1970), followed by the condensation reaction (3).

(1) $2\,(2\,HCOO^- + H^+ \rightarrow HCO_3^- + H_2CO)$
(2) $2\,H_2CO + OH^- \rightarrow CH_3OH + HCOO^-$
(3) $CH_3OH + HCOO^- \rightarrow CH_3COO^- + H_2O$

$4\,HCOO^- + H^+ \rightarrow 2HCO_3^- + CH_3COO^-$

Oxidation reactions are not thought to be a factor in these experiments: formic acid oxidation to CO_2 and H_2O is slow in the presence of pressurized oxygen, and undetectable at concentrations of 10^{-4}mol/L O_2. Formate oxidation to bicarbonate is approximately six times slower than formic acid oxidation (Bjerre and Sorensen, 1992).

Bjerre, A.B., Sorensen, *E., Ind. Eng. Chem. Res.* (1992) **31**, 1577-1580

Elliot, Douglas C., Hallen, Richard T., and Sealock Jr., L. John Ind. Eng. Chem. Prod. Res. Dev. (1983) **22**, 431-435

Matsuura T., Smith, J.M., *AICHE Journal* (1970) **16**, 1064-1071

Mayer, R., Runge, K., Dreschel, H*., Z. Chem* (1963) **3**, 134

Sinnarkar, N.D., Ray, M.N., *Indian J. Chem.* (1975) **13**, 962-964

Creating a Proto-Organism Through Lowest Entropy Departure Path

Steen Rasmussen, Liaohai Chen, Stirling Colgate, Klaus Lackner, and Martin Nilsson
[steen, chen, colgate, ksl, nilsson @lanl.gov]
Los Alamos National Laboratory, Los Alamos MN 87545, USA.

We seek to construct a biochemical system that can derive energy from a coupled redox complex—a simple form of metabolism—and use information carried in proto-genes where both complexes are integrated in a lipid aggregate. This proto-organism can self-replicate, use energy and nutrients available from its environment, undergo evolutionary change over time, and ultimately die. All the molecular species used are either simulated readily in prebiotic chemical reactions, or are present as organic components of carbonaceous chondrites and can thus be assumed to have existed on Earth some 4 billion years ago. Although we use such simple building blocks, we do not seek to reconstruct the origin of life on Earth. Indeed, our focus is on fundamental processes that could lead to the self-organization of matter into functional units and, thereby, helps describe the formation of life anywhere in the universe.

Two experimental systems have to be integrated in the laboratory and in simulation combining the three critical physico-chemical structures in Fig 1. A templating polymer (proto-gene) with a hydrophobic backbone performs template directed replication at a lipid interface (compartment). By using the slight energy advantage from the lipid-template complex compared to them being separated, more interface can form. A following hybridization reaction between the original template and two complementary oligomers is also thermodynamically down hill. As this new three component complex sinks into the hydrophobic environment, ligation (polymerization) becomes thermodynamically favorable as a means for generating more templates. Thus, a weak energetic coupling exists between templating and interface generation.

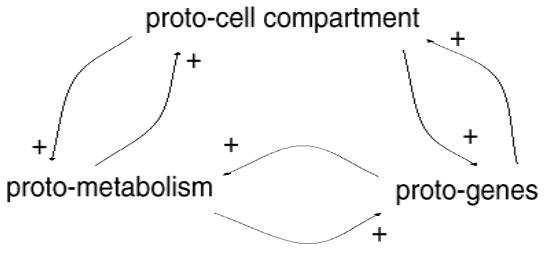


Fig 1. The proto-organism is a cooperation between the three molecular aggregates.

Since this system is only utilizing the slight free energy associated with lipid/template complexation and polymer templating, the process is very fragile. Using the chemical energy bound in mineral sulfur- or iron compounds (Fig 1, proto-metabolism), orders of magnitude in speed and precision can be obtained, and perhaps such a proto-metabolic system should be the starting point [Chen et al, *J.Phys. Chem. A* (1998) 102 (45), 9095-

9098]. A coupling between the production of e.g. PNA (peptide nucleic ascid) dimers and amphiphilic polymers can occur by designing an electron relay system in a vesicle membrane and using HS- as the energy source. This would constitute a continuous synthesis of surfactant and PNA dimer in a vesicle system. By functionalizing the PNA backbone certain PNA strands could enhance the redox reactions and an autocatalytic cooperation could be established between the templating and surfactant production processes. This feedback also forms the basis for a simple Darwinian evolution of the system (Rasmussen et al, 2001, *Artificial Life*, in press).

Although the proposed self-reproducing molecular aggregate does not constitute a contemporary cell, it is the first model of a concrete molecular system that results from thermodynamically downhill processes and that combines vesicles, metabolism, and proto-genes in a cooperative manner. We seek to address how the transition between nonliving and living matter relates to formal measures of complexity and what the entropy landscape is for this transition.

In parallel with the experimental system we are developing a computational version of the full proto-organism, using several different methods extending our current molecular simulation capabilities: Reaction kinetics-, MD-, MD lattice gas- [Nilsson et al, in *New Constructions in Cellular Automata*, D. Griffeath and C. Moore, eds, Oxford University Press (2001) 157-184 (in press); Mayer & Rasmussen, *Int. J. Modern Phys. C*, Vol 11, No. 4 (2000)], methods, although we can only investigate subsystems using the MD methods. Thus, we are in the process of developing a method to upscale and downscale in molecular systems ranging from Angstroms to microns and from the pico-second range to minutes.

Since our approach adopts a broad planetary perspective, we address questions regarding the formation of habitable planets and the signatures of possible life on these planets—what comes before and what comes after the transition from nonliving to living matter. If life is common in the universe, then planet formation must, at least some of the time, lead to the conditions necessary for life. Based on observations, theory, and simulations, we are currently trying to demonstrate that the accretion of carbonaceous chondrites in stellar nebulae could provide the "glue" that is missing in current accretionary theory in allowing the transition from dust-size grains to self-gravitating kilometer-size planetesimals. Thus, we propose to show that the same organic molecules that are necessary for life are also necessary for the accretion of habitable planets [Li, Colgate, Wendroff., Liska, *Astro. Phys.* J (2001) in press].

Our objective is to develop, comprehend, and test a novel paradigm for a universal, astrophysically- and cosmochemically-based theory of the origin of life. By demonstrating how to bridge nonliving and living matter in the laboratory, and by providing new methods for detecting signs of early and simple life on this planet, we can expect to extend this understanding to more distant worlds.

Evolution of Proteins

Monica Riley
Marine Biological Laboratory
Woods Hole, MA 02543
USA
mriley@mbl.edu

Tom McCormack
Marine Biological Laboratory
Woods Hole, MA 02543
USA

Laila Nahum
Marine Biological Laboratory
Woods Hole, MA 02543
USA

Alastair Kerr
Cereon, Corp
Boston, MA
USA
Alastair.Kerr@cereon.com

Ping Liang
Marine Biological Laboratory
Woods Hole, MA 02543
USA

The history of the evolution of proteins on planet earth can be reconstructed in important ways by the study of contemporary proteins. The exapansion from simple beginnings of kinds and capabilities of proteins is a history of gene duplication and divergence.

Families of contemporary proteins that seem likely to have shared ancestors have been assembled using information on protein sequences derived from fully sequenced microbial genomes. Clustering of related proteins based on sequence comparisons both within and between organisms has produced large classes of evolutionarily related enzymes, regulators and transporters. Using recursive sequence analysis, boundaries of groups of like proteins have been extended such that one can visualize the mechanisms by which ancestral proteins have diverged to yield progeny proteins with differences in function.

We have examined relationships among evolutionarily related enzymes of intermediary metabolism and have found that the chemistry of the reaction is conserved in most cases, specificity for ligands in some cases, and both attributes in many cases.

Going beyond sequence analysis, one can connect the amino acid sequence information with structural information, to see familial relationships based on structure that can no

longer be detected on the basis of sequence alone. In these ways it is becoming possible to group classes of proteins together that plausibly came from one or a few ancestral proteins. Examples will be given.

Thus we have found that it is possible to look far back in time to important stages of the origin of life, stages that preceded the existence of the Last Universal Common Ancestor.

Expanding the Limits of RNA Catalysis

Michael P. Robertson
Institute for Cellular and Molecular Biology
University of Texas at Austin
2500 Speedway, Austin, TX 78712
USA
mprobertson@mail.utexas.edu

Andrew D. Ellington
Institute for Cellular and Molecular Biology
University of Texas at Austin
USA

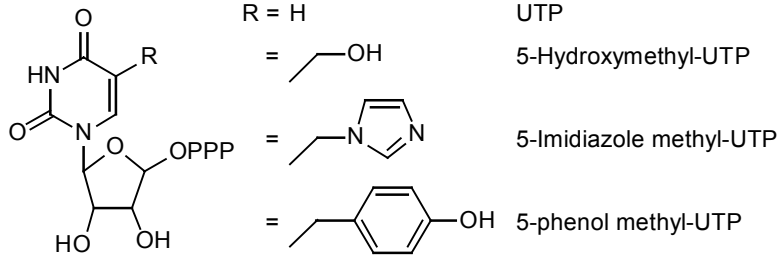
The RNA World hypothesis proposes that at some period during the evolution of life, RNA molecules performed the dual functions of genetic storage and enzymatic catalysis. The possibility that RNA can function in both these capacities was demonstrated with the discovery of catalytic RNA in the early 1980s. Since then the types of chemical reactions that RNA has been shown to catalyze has steadily increased and with it the feasibility of an all-RNA metabolism. However, a key aspect of modern biocatalysis is the ability of enzymes to attenuate their level of activity in response to changing environmental conditions. Presumably, a similar regulation strategy would also be necessary to sustain a reasonably complex RNA-based metabolism.

We have isolated and designed several RNA ligase ribozymes whose activity are regulated by the presence of a variety of different effectors. An oligonucleotide-dependent ligase (L1) was isolated from a randomized pool of RNA using in vitro selection. This ribozyme's activity in the presence of a specific oligonucleotide effector is 10,000-fold greater than when the effector is absent. Other allosterically regulated ribozymes were engineered by replacing a non-essential stem loop of L1 with various small molecule binding RNA aptamers to create aptamer-ribozyme hybrids (aptazymes). Ligand binding induces a conformational change in the aptamer domain which is propagated to the ribozyme domain through a short linking stem which allows the ribozyme domain to adopt an active conformation. The aptazyme's level of response to ligand can be adjusted by altering the basal stability of the stem linking the aptamer and ribozyme domains. This approach has resulted in aptazymes that are regulated by ATP, theophylline, and FMN with activations of 800, 1600, and 250 respectively.

Alternatively, we have employed a two stage in vitro selection procedure with a partially randomized pool based on the L1 ligase to isolate protein-activated aptazymes that are extremely dependent on their cognate effector for activity. Using this strategy, Cyt18 (a tyrosyl tRNA synthetase) and lysozyme dependent ligases have been isolated with activations of 75000 and 3000 respectively. The aptazyme's activation is not only specific for the protein it was selected against, but also recognizes only the native conformation of the protein.

In addition to the regulation of RNA catalysis, we have been investigating ways to increase the proficiency and versatility of ribozymes by augmenting the chemical functional groups of RNA by the incorporation of modified nucleosides. Three different uridine triphosphate (UTP) derivatives with modifications at the 5 position of the uracil ring (5-hydroxymethyl-, 5-imidazolemethyl-, and 5-phenolmethyl-UTP) which mimic the side chains of the amino acids serine, histidine, and tyrosine have been synthesized and tested (Figure 1). These modified nucleotides can be efficiently transcribed into RNA using T7 RNA polymerase and have been incorporated into random sequence RNA pools for use in *in vitro* selection experiments.

Figure 1



These results demonstrate that ribozyme activity can be regulated in much the same way that protein enzymes are regulated in contemporary biochemistry. This ability of RNA to be controlled by various potential metabolic intermediates contributes an additional layer of sophistication to ribozyme catalysis and increases the plausibility that a complex metabolism based solely on RNA could have once existed.

Peptide Formation in Hydrothermal Environments

David S. Ross, Kono Lemke
U.S. Geological Survey
345 Ravenswood Ave.
Menlo Park, CA
US
dsross3@yahoo.com

Koichiro Matsuno
Dept. of BioEngineering
Nagoaka University of Technology
Japan

The formation of prebiologically significant oligomers, including peptides and polynucleotides, is challenged by both thermodynamic and chemical kinetics factors, and the nature of their initial genesis is still not clear after more than 3 decades of study. For peptides it has been recognized that amide formation in aqueous solution is uphill by roughly 2 kcal/mol (25°C), and successive condensation will quickly lead to peptide concentrations well below meaningful values after just a few steps. Amide formation is estimated to be very slow, moreover, and even if there were there some means of eliminating reverse hydrolysis the estimated characteristic times of homogeneous condensation to dipeptides at 25°C and around pH 7 are about $5x10^{11}$ yr. The times fall to about $5x10^{7}$ yr only at temperatures above 100°C, but growth to larger oligomers would obviously be an unacceptably lengthy process even in astrobiological terms.

We present here our consideration of recent data in experiments simulating the oceanic circulation of water through hydrothermal zones (Imai, et al., 1999), where glycine was converted to a range of polyglycines at 225°C. The analysis shows that prebiotic peptide formation may not be as difficult as has been projected. While the equilibrium levels of diglycine observed were about a factor of 2 greater than predicted thermochemically, a deviation perhaps within error limits, we estimate that the equilibrium quantities of triglycine exceeded prediction by factor a factor of about 200. The larger peptides observed were present at even greater divergences from prediction.

The rates of peptide formation were similarly surprising: di- and triglycine attained equilibrium values over periods 30-40 times more rapidly than predictions based on known hydrolysis rate constants and the pertinent thermochemistry. Their rates of growth, moreover, were exponential, apparently reflecting the circulation of the reactant stream through temperature zones. Autocatalysis is known in some aqueous oxidation processes involving highly reactive free radical intermediates, but it has not been recognized in simple condensations such as these. These results suggest a means for peptide formation involving self-assisted condensation, a process which could develop through a kinetic and thermodynamic sequence stimulated by the thermal cycling of the reactant stream. Such a path could be vital to the ultimate formation of functional proteins in the prebiotic oceans.

A Chiroselective Peptide Replicator

Alan Saghatelian, Yohei Yokobayashi, Kathy Soltani, and M. Reza Ghadiri*
Departments of Chemistry and Molecular Biology and the Skaggs Institute for Chemical Biology
The Scripps Research Institute, La Jolla, CA 92037
USA
saghat@scripps.edu

The origin of homochirality is often addressed from the perspective of enantioselection that seeks to evoke various scenarios for the production of enantiomeric differences within the prebiotic chiral pool.[1-6] Although several plausible physicochemical processes have been suggested to be capable of producing minute enantiomeric differences, none can produce significant imbalances required for the production of homochiral biopolymers without some selectable chiral amplification process(es).[1,7-10] Therefore, it is believed that the origins of homochirality in living systems must have been based on selection and chiroselective amplification processes.[1,11-13] Yet despite its central importance, the feasibility of nonenzymatic template-directed chiroselective amplification in biopolymers had not been previously demonstrated.[13-16] Here we report that a 32-residue peptide replicator is capable of amplifying homochiral products efficiently and specifically through a chiroselective autocatalytic cycle. The chiroselective amplification is highly robust and can discriminate amongst structures possessing even single stereochemical mutations within otherwise homochiral sequences. Remarkably, the system also exhibits a hitherto unknown dynamic stereochemical editing function. The editing function is manifested through certain heterochiral sequences—mutants arising through uncatalyzed background fragment coupling reactions— that posses potent cross-catalytic activities for promoting the production of the homochiral product. This study thus lends support to the postulate that self-replicating polypeptides could have played a role in the origin of homochirality on Earth.

References:

1. For a review see: Bonner, W. A. The origin and amplification of biomolecular chirality. *Origins Life Evol. Biosphere* **21**, 59-111 (1991).
2. Bonner, W. A., Blair, N. E. & Dirbas, F. M. Experiments on the abiotic amplification of optical activity. *Origins Life* **11**, 119-134 (1981).
3. Cronin, J. R. & Pizzarello, S. Enantiomeric excesses in meteoritic amino acids. *Science (Washington, D. C.)* **275**, 951-955 (1997).
4. Kondepudi, D. K. & Nelson, G. W. Weak neutral currents and the origin of biomolecular chirality. *Nature (London)* **314**, 438-441 (1985).
5. Kondepudi, D. K., Kaufman, R. J. & Singh, N. Chiral symmetry breaking in sodium chlorate crystallization. *Science (Washington, D. C., 1883-)* **250**, 975-977 (1990).
6. Weissbuch, I. *et al.* Spontaneous generation and amplification of optical activity in .alpha.-amino acids by enantioselective occlusion into centrosymmetric crystals of glycine. *Nature (London)* **310**, 161-164 (1984).

7. Frank, F. C. On spontaneous asymmetric synthesis. *Boichim. Biophys. Acta* **11**, 459-463 (1953).
8. Calvin, M. *Chemical evolution* (Oxford University Press, Oxford, 1969).
9. Girard, C. & Kagan, H. B. Nonlinear effects in asymmetric synthesis and stereoselective reactions: ten years of investigation. *Angew. Chem., Int. Ed.* **37**, 2923-2959 (1998).
10. Soai, K., Shibata, T., Morioka, H. & Choji, K. Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. *Nature (London)* **378**, 767-768 (1995).
11. Siegel, J. S. Homochiral imperative of molecular evolution. *Chirality* **10**, 24-27 (1998).
12. Bada, J. L. Origins of homochirality. *Nature (London)* **374**, 594-595 (1995).
13. Bolli, M., Micura, R. & Eschenmoser, A. Pyranosyl-RNA: chiroselective self-assembly of base sequences by ligative oligomerization of tetranucleotide-2',3'-cyclophosphates (with a commentary concerning the origin of biomolecular homochirality). *Chem. Biol.* **4**, 309-320 (1997).
14. Joyce, G. F. *et al.* Chiral selection in poly(C)-directed synthesis of oligo(G). *Nature (London)* **310**, 602-604 (1984).
15. Schmidt, J. G., Nielsen, P. E. & Orgel, L. E. Enantiomeric Cross-Inhibition in the Synthesis of Oligonucleotides on a Nonchiral Template. *J. Am. Chem. Soc.* **119**, 1494-1495 (1997).
16. Kozlov, I. A., Politis, P. K., Pitsch, S., Herdewijn, P. & Orgel, L. E. A Highly Enantio-Selective Hexitol Nucleic Acid Template for Nonenzymic Oligoguanylate Synthesis. *J. Am. Chem. Soc.* **121**, 1108-1109 (1999).

When Can You Attribute Biological Origins to Small Organic Molecules?

Roger E. Summons
Australian Geological Survey Organisation
PO Box 378, Canberra, ACT 2601
Australia
roger.summons@agso.gov.au

The carbon chemistry of living systems is one of the most exactly known branches of all science and this chemistry provides precise tools for identifying and characterising extant and extinct life on Earth and, potentially, throughout the universe. Lipids are especially ubiquitous and recalcitrant small organic compounds that are common to all living things and which carry structural and isotopic information pertaining to their mode of synthesis. Hydrocarbons, which encode diagnostic structural and isotopic information, are the most readily preserved lipid end-product and are widely distributed in terrestrial sediments of all ages. Hydrocarbons are also abundant components of carbonaceous meteorites. This paper summarises deliberations of a group[1] from the NASA Task Force on Biosignatures for Mars Exploration and presents some criteria that assist in distinguishing between biogenic and abiogenic hydrocarbons. Characteristics of biogenic hydrocarbons include:

- Enantiomeric excess
- Preference for certain diasteromers
- Constitutional isomer preference
- Repeating constitutional sub-units or atomic ratios
- Clusters or patterns in carbon number, concentration for related compounds
- Systematic isotopic ordering at molecular level

Their utility can be tested through comparisons of hydrocarbon distributions for ancient terrestrial sediments with those reported for meteorites.

[1]Roger E. Summons, Pierre Albrecht, Sherwood Chang, Gene McDonald, and J. Michael Moldowan (2001) Molecular Biosignatures. In: J.F. Kerridge (ed.) unpublished report of NASA Task Force on Biosignatures for Mars Exploration.

Alternative Nucleic Acids

Christopher Switzer
Chemistry Department
University of California, Riverside
Riverside, CA 92521
USA
christopher.switzer@ucr.edu

John C. Chaput
Azizul Haque
Dong-Kye Lee
Paul H. Kim
Surajit Sinha
Chemistry Department
University of California, Riverside
Riverside, CA 92521
USA

Our goal is to synthesize Alternative Nucleic Acids (ANAs) to attempt the optimization of polymer structure subject to the constraints of prebiotic availability, template-directed reproduction, replication conservative mutation, and fitness. We have identified ANAs by taking small steps in "structure-space" away from RNA (the best model for a molecule bearing features both universal and unique to life) that may avoid some of the problems inherent in fulfillment of the aforementioned constraints. ANAs are under investigation with novel perturbations to (i) base-pairing domains (ii) formal charges, and (iii) the phosphodiester backbone. These studies help to define chemical parameters for molecular evolution. The work also addresses whether nucleic acid-like molecules are sufficient to enable the origin of life and what limitations exist for life elsewhere in the universe based on a single biopolymer (eg. RNA) rather than multiple biopolymers (DNA, RNA, proteins, carbohydrates). Recent accomplishments in these areas will be presented.

Evolution of Gene Order in Prokaryotes

Javier Tamames
Centro de Astrobiología (CAB)
INTA/CSIC
Carretera de Ajalvir Km.4, Torrejón de Ardoz, Madrid
SPAIN
tamames@cnb.uam.es

María González
Centro de Astrobiología (CAB)
INTA/CSIC
SPAIN

Alfonso Valencia
Centro de Astrobiología (CAB)
INTA/CSIC
SPAIN

Gene order conservation (GOC) has been proposed as a very useful tool for predicting interactions between the proteins encoded in these genes (Huynen et al, 2000). Also it has been proposed as a tool for studying the evolution of genomes and organisms (Lathe et al, 2000).

We have performed an extensive study of GOC between 32 bacterial and 7 archaeal genomes. Our procedure runs alignments between all ORFS for every pair of genomes, and identifies orthologues and clusters with gene order conservation. For studying the relation between GOC and evolution, distances between organisms are estimated using both classical phylogenies using SSU rRNA sequences and genomic methods such as common gene content between species.

The results show that GOC is a labile characteristic of genomes, that are considerably rearranged during evolution. Nevertheless, GOC is still detectable at medium phylogenetic distances (those between E.coli and B.subtilis, for instance), and therefore can be valuable information for studying evolution. GOC is a better measure of relationships between organisms than common gene content, and can help to uncover and solve inconsistencies in classical phylogenies. It can be used also for studying important evolutionary traits such as lateral gene transfer between organisms.

Even if GOC tends to be lost during evolution, some specially conserved clusters are apparent. The cluster of ribosomal proteins is the best known example. This conservation implies an active selection process mantaining the structure of the clusters. Recently we have published new hypothesis for this selection. At least in one case (cluster dcw taking part in cell wall byosynthesis and cell division), it is likely the existence of a link between the organization of the cluster and phenotipic properties, celullar morphology in this case

(Tamames et al, 2001). This is in accordance with previous hypothesis relating morphology of the bacterial cells with evolution (Seifert and Fox, 1998).

References

Huynen M, Snel B, Lathe W, Bork P. (2000) *Genome Res* 10, 1204-10

Lathe, W., Snel, B., Bork, P. (2000) *Trends Biochem Sci* 25, 474-479.

Siefert, J., Fox, G. (1998) *Microbiology* 144, 2803-2808.

Tamames, J., González, M., Mingorance, J., Valencia, A., Vicente, M. (2001) *Trends Genet.*, In the press

Experimental Paleobiochemistry: Understanding Major Transitions in Life on Earth

J. Michael Thomson
Dept. Of Anatomy and Cell Biology
University of Florida
1600 SW Archer Rd. Rm. B1-3
Gainesville, FL. 32610-0235
USA
mthomson@grove.ufl.edu

John P. Aris
Dept of Anatomy and Cell Biology
University of Florida
USA

Steven A. Benner
Dept of Anatomy and cell Biology
University of Florida
USA

Astrobiology requires integrating concepts from natural history with concepts from the physical sciences. The former includes geology, paleontology, and paleoecology; the latter includes chemistry, physics, and molecular biology. In the past, this integration has required collaborations between individuals and laboratories in different disciplines. In the future, a single individual working in a single laboratory will be trained at this level of "comprehensivity", and to design insightful research projects.

Following this strategy, we have explored the interaction between fungi and their environment over the past 100 million years, a period over which the planet has undergone dramatic upheavals in the climate that it carries and the life that it supports. We have reconstructed elements of the molecular biology of the yeasts that invaded the fruit, and gave rise to the modern process known as "fermentation". These fungi emerged to exploit the new resource made available by angiosperms (fruiting plants), that began to provide the dominant floral genera over much of the planet towards the end of the Cretaceous. Resurrection of these ancestral proteins in a "paleobiochemistry" experiment, and data collected on their physical chemical behavior in the laboratory, provide insight into how this metabolic process arose through the interaction between the fungus and a host in the paleoecosystem. Further, experimental paleobiochemistry permits us to understand features of the "downstream" paleoecology, including the interaction of fruit flies and other frugivorous animals with fruit in the ecology that emerged following the extinction of the dinosaurs. What follows is a remarkably complete understanding of "function", extending from the molecules to the planet.

We have also focused on the reliability of ancestral reconstructions inferred by the analysis of DNA and protein sequences from contemporary organisms. Enhanced reliability comes, of course, from larger collections of descendent sequences, and we

have collected 15 of these from a collection of closely related yeasts. The robustness of the reconstructions with respect to different methods for their creation has also been explored. We have also developed tools for dating molecular changes with respect to altered developmental function. These changes accompany those of the surrounding environmental constraints placed on the interactions between fungi, fruit, and fruit flies, and extend a better understanding of the evolution of early life on earth.

De Novo Catalysts of Biopolymer Synthesis

Keith M. Wilcoxen
Department of Chemistry
The Scripps Research Institute
10660 Torrey Pines Rd, San Diego, CA 92122
USA
wilcoxen@scripps.edu

Dana Weinberg
Department of Chemistry
The Scripps Research Institute
10660 Torrey Pines Rd, San Diego, CA 92122
USA
danaw@scripps.edu

M. Reza Ghadiri
Department of Chemistry
The Scripps Research Institute
10660 Torrey Pines Rd, San Diego, CA 92122
USA
ghadiri@scripps.edu

Nucleophilic catalysis of acyl transfer reactions is a fundamental theme in the biological synthesis of many naturally occurring biopolymers and natural products. The ability of enzymes to carry out selective acyl transfer reactions has given the cellular organism a vast array of molecules needed for sustained life. Simple carbonyl reactions have provided a biological basis for carbon-carbon bond formation necessary for the production of peptides, fatty acids, polyesters, polynucleic acids, polysaccharides and most secondary metabolites, many of which are formed by the catalytic properties of another biopolymer. The formation of many of these biopolymers is achieved by the repetitive and controlled living polymerization of thioester monomers covalently bound to condensing enzymes through a thioester bond. Among these, catalytic amide bond formation has proven to be a versatile and ubiquitous method for the biosynthesis of proteins and non-ribosomal peptides. It is under this context that we have selected the sequence of a 33-residue α-helical peptide, GCN4-LI, which is known to assemble into a stable, homomeric 4-helix bundle, as the basis scaffold for our acyl transfer catalysts, with catalytic residues located near the interface of the helices in the helical bundle. We wish to report initial results toward the development of *de novo* scaffold and active site engineering of an amino acid condensing enzyme, mimicking the properties of non-ribosomal peptide *synthase*, in an effort to achieve elementary catalytic function in a versatile and programmable system.

Mosaicism in 16S rRNA Genes

Olga A. Zhaxybayeva, Lorraine Olendzenski, J. Peter Gogarten
Department of Molecular and Cell Biology
University of Connecticut
U-3044, Storrs, CT, 06269
USA
olga.zh@uconn.edu

The observation that genome content trees and traditional small subunit ribosomal RNA phylogenies agree with each other has been used to discount the importance of horizontal gene transfer (HGT) (e.g., Snel et al., 1999; Tekaia et al., 1999). However, this congruence is also in perfect agreement with the theory that prokaryotic taxonomic units might exclusively reflect HGT frequency and not vertical inheritance: If two organisms appear related because they more frequently exchange genes with each other; their genome content becomes more similar, and at the same time each individual gene tree becomes more likely to group the two particular organisms together, the more frequently the organisms exchanged genes. Nevertheless, other gene trees often show strong well supported conflicts with 16S and with genome content trees (e.g. Klenk et al., 1999; Olendzenski et al., 2000), thus the congruence between genome content and 16S rRNA trees remains somewhat surprising.

It has been argued that rRNA is such a good marker for organismal evolution, because it interacts with so many cellular components, and therefore cannot readily be transferred between organisms. On the other hand, many antibiotics are targeted against the translation machinery providing a strong selective advantage to those organisms that can evade the antibiotic's action. Furthermore, it has been known for a long time that functioning ribosomes can be reconstituted from parts taken from different organisms, and the ribosomal operons of an organism can be replaced under laboratory conditions and *in vivo* with rRNA operons from another species (e.g. Nomura et al., 1968; Bellemare et al., 1973; Wrede et al., 1973; Yap et al., 1999; Asai et al., 1999).

The potential difference between a molecular or gene tree and organismal evolution has been widely recognized. The latter is often assumed to be net-like or reticulate (e.g.: Hilario et al., 1993; Gogarten, 1995). However, genes are not immutable units of inheritance, and even at the gene level evolution is not always treelike. It is known that if multiple identical or similar copies of a gene are present in the same cell, gene conversion events tend to make these two copies of the genes more similar to one another (cf. Gogarten et al., 1999). These conversion tracts, i.e., the stretches that are copied from one gene to another, are usually in the range of only a few hundred nucleotides (Sweetser et al., 1994; Betran et al., 1997; Yang et al., 1997) - much smaller than the lengths of a typical gene. The result is that genes do not evolve as a whole, rather different parts of the genes can have different histories.

rRNA is very conserved at the nucleotide level providing many recombination points inside the RNA gene itself. The mosaic nature of rRNA operons can explain the

congruence of the rRNA data with gene content trees: The rRNA encoding genes are mosaic themselves, reflecting the mosaic character of the genome as a whole. To test the possibility that the rRNA itself is a mosaic, we retrieved 100 aligned representative bacterial 16rRNA sequences from the RDP database (http://www.cme.msu.edu/RDP/html/index.html). The alignment was divided into a number of small fragments of arbitrary length. The full-length alignment and individual parts of the alignment were analyzed using neighbor joining method under different distance measures. 100 bootstrap replicates were generated for every data set. Each tree topology obtained from a fragment was compared to the tree topology reconstructed from full-length alignment. Only conflicts with bootstrap support above 70 were considered.

We found seventeen conflicts between topology reconstructed from full-length alignment and topologies based on different fragments. The disagreements persisted when we repeated the analyses using parsimony method as implemented in PAUP* (Swofford, 1998). For each conflict additional extended datasets were constructed that contained many additional sequences more closely related to the conflicting taxa. Analysis of these datasets showed that the detected conflicts remained in the presence of additional sequences for most of the seventeen cases.

Our findings indicate that rRNA genes were frequently transferred between related organisms, and that the transferred genes recombined with those already present in organisms. This finding is in agreement with the sizes of DNA fragments that are usually transferred between microbes, with reports of transfer of rRNA genes and ribosomal operons, and with concepts that suggest that horizontal gene transfer frequency might be one of the main determinants of taxonomic units among prokaryotes.

References

Asai, T., D. Zaporojets, C. Squires and C. L. Squires (1999) PNAS **96**(5): 1971-6.
Bellemare, G., R. Vigne and B. R. Jordan (1973) Biochimie **55**(1): 29-35.
Betran, E., J. Rozas, A. Navarro and A. Barbadilla (1997) Genetics **146**(1): 89-99.
Gogarten, J. P. (1995) Trends in Ecology and Evolution **10**: 147-151.
Gogarten, J. P. and L. Olendzenski (1999) Curr Opin Genet Dev 9(6): 630-6.
Hilario, E. and J. P. Gogarten (1993) Biosystems **31**(2-3): 111-9.
Klenk, H. P., T. D. Meier, P. Durovic, V. Schwass, F. Lottspeich, P. P. Dennis and W. Zillig (1999) J Mol Evol **48**(5): 528-41.
Nomura, M., P. Traub and H. Bechmann (1968) Nature **219**(156): 793-9.
Olendzenski, L., L. Liu, O. Zhaxybayeva, R. Murphey, D. G. Shin and J. P. Gogarten (2000) J Mol Evol **51**(6): 587-99.
Snel, B., P. Bork and M. A. Huynen (1999) Nat Genet **21**(1): 108-10.
Sweetser, D. B., H. Hough, J. F. Whelden, M. Arbuckle and J. A. Nickoloff (1994) Mol Cell Biol **1**4(6): 3863-75.
Swofford, D. (1998). PAUP* 4.0 , Sinauer Associates, Inc., Sunderland, MA, USA.
Tekaia, F., A. Lazcano and B. Dujon (1999) Genome Res 9(6): 550-7
Wrede, P. and V. A. Erdmann (1973) FEBS Lett **33**(3): 315-9.
Yang, D. and A. S. Waldman (1997) Mol Cell Biol **17**(7): 3614-28.
Yap, W. H., Z. Zhang and Y. Wang (1999). J Bacteriol **181**(17): 5201-9.